Mouse Dnmt3a DNA sequence

```
   1 GAATTCCGGC CTGCTGCCGG GCCGCCCGAC CCGCCGGGCC ACACGGCAGA
  51 GCCGCCTGAA GCCCAGCGCT GAGGCTGCAC TTTTCCGAGG GCTTGACATC
 101 AGGGTCTATG TTTAAGTCTT AGCTCTTGCT TACAAAGACC ACGGCAATTC
 151 CTTCTCTGAA GCCCTCGCAG CCCCACAGCG CCCTCGCAGC CCCAGCCTGC
 201 CGCCTACTGC CCAGCAATGC CCTCCAGCGG CCCCGGGGAC ACCAGCAGCT
 251 CCTCTCTGGA GCGGGAGGAT GATCGAAAGG AAGGAGAGGA ACAGGAGGAG
 301 AACCGTGGCA AGGAAGAGCG CCAGGAGCCC AGCGCCACGG CCCGGAAGGT
 351 GGGGAGGCCT GGCCCGAAGC GCAAGCACCC ACCGGTGGAA AGCAGTGACA
 401 CCCCCAAGGA CCCAGCAGTG ACCACCAAGT CTCAGCCCAT GGCCCAGGAC
 451 TCTGGCCCCT CAGATCTGCT ACCCAATGGA GACTTGGAGA AGCGGAGTGA
 501 ACCCCAACCT GAGGAGGGGA GCCCAGCTGC AGGGCAGAAG GGTGGGGCCC
 551 CAGCTGAAGG AGAGGGAACT GAGACCCCAC CAGAAGCCTC CAGAGCTGTG
 601 GAGAATGGCT GCTGTGTGAC CAAGGAAGGC CGTGGAGCCT CTGCAGGAGA
 651 GGGCAAAGAA CAGAAGCAGA CCAACATCGA ATCCATGAAA ATGGAGGGCT
 701 CCCGGGGCCG ACTGCGAGGT GGCTTGGGCT GGGAGTCCAG CCTCCGTCAG
 751 CGACCCATGC CAAGACTCAC CTTCCAGGCA GGGGACCCCT ACTACATCAG
 801 CAAACGGAAA CGGGATGAGT GGCTGGCACG TTGGAAAAGG GAGGCTGAGA
 851 AGAAAGCCAA GGTAATTGCA GTAATGAATG CTGTGGAAGA GAACCAGGCC
 901 TCTGGAGAGT CTCAGAAGGT GGAGGAGGCC AGCCCTCCTG CTGTGCAGCA
 951 GCCCACCGAC CCTGCTTCTC CGACTGTGGC CACCACCCCT GAGCCAGTAG
1001 GAGGGGATGC TGGGGACAAG AATGCTACCA AAGCAGCCGA CGATGAGCCT
1051 GAGTATGAGG ATGGCCGGGG CTTTGGCATT GGAGAGCTGG TGTGGGGGAA
1101 ACTTCGGGGC TTCTCCTGGT GGCCAGGCCG AATTGTGTCT TGGTGGATGA
```

FIG. 1A-1

1151 CAGGCCGGAG CCGAGCAGCT GAAGGCACTC GCTGGGTCAT GTGGTTCGGA

1201 GATGGCAAGT TCTCAGTGGT GTGTGTGGAG AAGCTCATGC CGCTGAGCTC

1251 CTTCTGCAGT GCATTCCACC AGGCCACCTA CAACAAGCAG CCCATGTACC

1301 GCAAAGCCAT CTACGAAGTC CTCCAGGTGG CCAGCAGCCG TGCCGGGAAG

1351 CTGTTTCCAG CTTGCCATGA CAGTGATGAA AGTGACAGTG GCAAGGCTGT

1401 GGAAGTGCAG AACAAGCAGA TGATTGAATG GGCCCTCGGT GGCTTCCAGC

1451 CCTCGGGTCC TAAGGGCCTG GAGCCACCAG AAGAAGAGAA GAATCCTTAC

1501 AAGGAAGTTT ACACCGACAT GTGGGTGGAG CCTGAAGCAG CTGCTTACGC

1551 CCCACCCCCA CCAGCCAAGA AACCCAGAAA GAGCACAACA GAGAAACCTA

1601 AGGTCAAGGA GATCATTGAT GAGCGCACAA GCGAGCGGCT GGTGTATGAG

1651 GTGCGCCAGA AGTGCAGAAA CATCGAGGAC ATTTGTATCT CATGTGGGAG

1701 CCTCAATGTC ACCCTGGAGC ACCCACTCTT CATTGGAGGC ATGTGCCAGA

1751 ACTGTAAGAA CTGCTTCTTG GAGTGTGCTT ACCAGTATGA CGACGATGGG

1801 TACCAGTCCT ATTGCACCAT CTGCTGTGGG GGGCGTGAAG TGCTCATGTG

1851 TGGGAACAAC AACTGCTGCA GGTGCTTTTG TGTCGAGTGT GTGGATCTCT

1901 TGGTGGGGCC AGGAGCTGCT CAGGCAGCCA TTAAGGAAGA CCCCTGGAAC

1951 TGCTACATGT GCGGGCATAA GGGCACCTAT GGGCTGCTGC GAAGACGGGA

2001 AGACTGGCCT TCTCGACTCC AGATGTTCTT TGCCAATAAC CATGACCAGG

2051 AATTTGACCC CCCAAAGGTT TACCCACCTG TGCCAGCTGA GAAGAGGAAG

2101 CCCATCCGCG TGCTGTCTCT CTTTGATGGG ATTGCTACAG GGCTCCTGGT

2151 GCTGAAGGAC CTGGGCATCC AAGTGGACCG CTACATTGCC TCCGAGGTGT

2201 GTGAGGACTC CATCACGGTG GGCATGGTGC GGCACCAGGG AAAGATCATG

2251 TACGTCGGGG ACGTCCGCAG CGTCACACAG AAGCATATCC AGGAGTGGGG

2301 CCCATTCGAC CTGGTGATTG GAGGCAGTCC CTGCAATGAC CTCTCCATTG

FIG. 1A-2

2351 TCAACCCTGC CCGCAAGGGA CTTTATGAGG GTACTGGCCG CCTCTTCTTT

2401 GAGTTCTACC GCCTCCTGCA TGATGCGCGG CCCAAGGAGG GAGATGATCG

2451 CCCCTTCTTC TGGCTCTTTG AGAATGTGGT GGCCATGGGC GTTAGTGACA

2501 AGAGGGACAT CTCGCGATTT CTTGAGTCTA ACCCCGTGAT GATTGACGCC

2551 AAAGAAGTGT CTGCTGCACA CAGGGCCCGT TACTTCTGGG GTAACCTTCC

2601 TGGCATGAAC AGGCCTTTGG CATCCACTGT GAATGATAAG CTGGAGCTGC

2651 AAGAGTGTCT GGAGCACGGC AGAATAGCCA AGTTCAGCAA AGTGAGGACC

2701 ATTACCACCA GGTCAAACTC TATAAAGCAG GGCAAAGACC AGCATTTCCC

2751 CGTCTTCATG AACGAGAAGG AGGACATCCT GTGGTGCACT GAAATGGAAA

2801 GGGTGTTTGG CTTCCCCGTC CACTACACAG ACGTCTCCAA CATGAGCCGC

2851 TTGGCGAGGC AGAGACTGCT GGGCCGATCG TGGAGCGTGC CGGTCATCCG

2901 CCACCTCTTC GCTCCGCTGA AGGAATATTT TGCTTGTGTG TAAGGGACAT

2951 GGGGGCAAAC TGAAGTAGTG ATGATAAAAA AGTTAAACAA ACAAACAAAC

3001 AAAAAACAAA ACAAAACAAT AAAACACCAA GAACGAGAGG ACGGAGAAAA

3051 GTTCAGCACC CAGAAGAGAA AAAGGAATTT AAAGCAAACC ACAGAGGAGG

3101 AAAACGCCGG AGGGCTTGGC CTTGCAAAAG GGTTGGACAT CATCTCCTGA

3151 GTTTTCAATG TTAACCTTCA GTCCTATCTA AAAAGCAAAA TAGGCCCCTC

3201 CCCTTCTTCC CCTCCGGTCC TAGGAGGCGA ACTTTTTGTT TTCTACTCTT

3251 TTTCAGAGGG GTTTTCTGTT TGTTTGGGTT TTTGTTTCTT GCTGTGACTG

3301 AAACAAGAGA GTTATTGCAG CAAAATCAGT AACAACAAAA AGTAGAAATG

3351 CCTTGGAGAG GAAAGGGAGA GAGGGAAAAT TCTATAAAAA CTTAAAATAT

3401 TGGTTTTTTT TTTTTTTCCT TTTCTATATA TCTCTTTGGT TGTCTCTAGC

3451 CTGATCAGAT AGGAGCACAA ACAGGAAGAG AATAGAGACC CTCGGAGGCA

3501 GAGTCTCCTC TCCCACCCCC CGAGCAGTCT CAACAGCACC ATTCCTGGTC

FIG. 1A-3

3551 ATGCAAAACA GAACCCAACT AGCAGCAGGG CGCTGAGAGA ACACCACACC

3601 AGACACTTTC TACAGTATTT CAGGTGCCTA CCACACAGGA AACCTTGAAG

3651 AAAACCAGTT TCTAGAAGCC GCTGTTACCT CTTGTTTACA GTTTATATAT

3701 ATATGATAGA TATGAGATAT ATATATATAA AAGGTACTGT TAACTACTGT

3751 ACATCCCGAC TTCATAATGG TGCTTTCAAA ACAGCGAGAT GAGCAAAGAC

3801 ATCAGCTTCC GCCTGGCCCT CTGTGCAAAG GGTTTCAGCC CAGGATGGGG

3851 AGAGGGGAGC AGCTGGAGGG GGTTTTAACA AACTGAAGGA TGACCCATAT

3901 CACCCCCCAC CCCTGCCCCA TGCCTAGCTT CACCTGCCAA AAAGGGGCTC

3951 AGCTGAGGTG GTCGGACCCT GGGGAAGCTG AGTGTGGAAT TTATCCAGAC

4001 TCGCGTGCAA TAACCTTAGA ATATGAATCT AAAATGACTG CCTCAGAAAA

4051 ATGGCTTGAG AAAACATTGT CCCTGATTTT GAATTCGTCA GCCACGTTGA

4101 AGGCCCCTTG TGGGATCAGA AATATTCCAG AGTGAGGGAA AGTGACCCGC

4151 CATTAACCCC NCCTGGAGCA AATAAAAAAA CATACAAAAT GT

FIG. 1A-4

Mouse Dnmt3b1 DNA Sequence

```
   1  GAATTCCGGG CGCCGGGGTT AAGCGGCCCA AGTAAACGTA GCGCAGCGAT
  51  CGGCGCCGGA GATTCGCGAA CCCGACACTC CGCGCCGCCC GCCGGCCAGG
 101  ACCCGCGGCG CGATCGCGGC GCCGCGCTAC AGCCAGCCTC ACGACAGGCC
 151  CGCTGAGGCT TGTGCCAGAC CTTGGAAACC TCAGGTATAT ACCTTTCCAG
 201  ACGCGGGATC TCCCCTCCCC CATCCATAGT GCCTTGGGAC CAAATCCAGG
 251  GCCTTCTTTC AGGAAACAAT GAAGGGAGAC AGCAGACATC TGAATGAAGA
 301  AGAGGGTGCC AGCGGGTATG AGGAGTGCAT TATCGTTAAT GGGAACTTCA
 351  GTGACCAGTC CTCAGACACG AAGGATGCTC CCTCACCCCC AGTCTTGGAG
 401  GCAATCTGCA CAGAGCCAGT CTGCACACCA GAGACCAGAG GCCGCAGGTC
 451  AAGCTCCCGG CTGTCTAAGA GGGAGGTCTC CAGCCTTCTG AATTACACGC
 501  AGGACATGAC AGGAGATGGA GACAGAGATG ATGAAGTAGA TGATGGGAAT
 551  GGCTCTGATA TTCTAATGCC AAAGCTCACC CGTGAGACCA AGGACACCAG
 601  GACGCGCTCT GAAAGCCCGG CTGTCCGAAC CCGACATAGC AATGGGACCT
 651  CCAGCTTGGA GAGGCAAAGA GCCTCCCCCA GAATCACCCG AGGTCGGCAG
 701  GGCCGCCACC ATGTGCAGGA GTACCCTGTG GAGTTTCCGG CTACCAGGTC
 751  TCGGAGACGT CGAGCATCGT CTTCAGCAAG CACGCCATGG TCATCCCCTG
 801  CCAGCGTCGA CTTCATGGAA GAAGTGACAC CTAAGAGCGT CAGTACCCCA
 851  TCAGTTGACT TGAGCCAGGA TGGAGATCAG GAGGGTATGG ATACCACACA
 901  GGTGGATGCA GAGAGCAGAG ATGGAGACAG CACAGAGTAT CAGGATGATA
 951  AAGAGTTTGG AATAGGTGAC CTCGTGTGGG GAAAGATCAA GGGCTTCTCC
1001  TGGTGGCCTG CCATGGTGGT GTCCTGGAAA GCCACCTCCA AGCGACAGGC
```

FIG. 1B-1

1051 CATGCCCGGA ATGCGCTGGG TACAGTGGTT TGGTGATGGC AAGTTTTCTG

1101 AGATCTCTGC TGACAAACTG GTGGCTCTGG GGCTGTTCAG CCAGCACTTT

1151 AATCTGGCTA CCTTCAATAA GCTGGTTTCT TATAGGAAGG CCATGTACCA

1201 CACTCTGGAG AAAGCCAGGG TTCGAGCTGG CAAGACCTTC TCCAGCAGTC

1251 CTGGAGAGTC ACTGGAGGAC CAGCTGAAGC CCATGCTGGA GTGGGCCCAC

1301 GGTGGCTTCA AGCCTACTGG GATCGAGGGC CTCAAACCCA ACAAGAAGCA

1351 ACCAGTGGTT AATAAGTCGA AGGTGCGTCG TTCAGACAGT AGGAACTTAG

1401 AACCCAGGAG ACGCGAGAAC AAAAGTCGAA GACGCACAAC CAATGACTCT

1451 GCTGCTTCTG AGTCCCCCCC ACCCAAGCGC CTCAAGACAA ATAGCTATGG

1501 CGGGAAGGAC CGAGGGGAGG ATGAGGAGAG CCGAGAACGG ATGGCTTCTG

1551 AAGTCACCAA CAACAAGGGC AATCTGGAAG ACCGCTGTTT GTCCTGTGGA

1601 AAGAAGAACC CTGTGTCCTT CCACCCCCTC TTTGAGGGTG GGCTCTGTCA

1651 GAGTTGCCGG GATCGCTTCC TAGAGCTCTT CTACATGTAT GATGAGGACG

1701 GCTATCAGTC CTACTGCACC GTGTGCTGTG AGGGCCGTGA ACTGCTGCTG

1751 TGCAGTAACA CAAGCTGCTG CAGATGCTTC TGTGTGGAGT GTCTGGAGGT

1801 GCTGGTGGGC GCAGGCACAG CTGAGGATGC CAAGCTGCAG GAACCCTGGA

1851 GCTGCTATAT GTGCCTCCCT CAGCGCTGCC ATGGGGTCCT CCGACGCAGG

1901 AAAGATTGGA ACATGCGCCT GCAAGACTTC TTCACTACTG ATCCTGACCT

1951 GGAAGAATTT GAGCCACCCA AGTTGTACCC AGCAATTCCT GCAGCCAAAA

2001 GGAGGCCCAT TAGAGTCCTG TCTCTGTTTG ATGGAATTGC AACGGGGTAC

2051 TTGGTGCTCA AGGAGTTGGG TATTAAAGTG GAAAAGTACA TTGCCTCCGA

2101 AGTCTGTGCA GAGTCCATCG CTGTGGGAAC TGTTAAGCAT GAAGGCCAGA

2151 TCAAATATGT CAATGACGTC CGGAAAATCA CCAAGAAAAA TATTGAAGAG

2201 TGGGGCCCGT TCGACTTGGT GATTGGTGGA AGCCCATGCA ATGATCTCTC

FIG. 1B-2

2251 TAACGTCAAT CCTGCCCGCA AAGGTTTATA TGAGGGCACA GGAAGGCTCT

2301 TCTTCGAGTT TTACCACTTG CTGAATTATA CCCGCCCCAA GGAGGGCGAC

2351 AACCGTCCAT TCTTCTGGAT GTTCGAGAAT GTTGTGGCCA TGAAAGTGAA

2401 TGACAAGAAA GACATCTCAA GATTCCTGGC ATGTAACCCA GTGATGATCG

2451 ATGCCATCAA GGTGTCTGCT GCTCACAGGG CCCGGTACTT CTGGGGTAAC

2501 CTACCCGGAA TGAACAGGCC CGTGATGGCT TCAAAGAATG ATAAGCTCGA

2551 GCTGCAGGAC TGCCTGGAGT TCAGTAGGAC AGCAAAGTTA AAGAAAGTGC

2601 AGACAATAAC CACCAAGTCG AACTCCATCA GACAGGGCAA AAACCAGCTT

2651 TTCCCTGTAG TCATGAATGG CAAGGACGAC GTTTTGTGGT GCACTGAGCT

2701 CGAAAGGATC TTCGGCTTCC CTGCTCACTA CACGGACGTG TCCAACATGG

2751 GCCGCGGCGC CCGTCAGAAG CTGCTGGGCA GGTCCTGGAG TGTACCGGTC

2801 ATCAGACACC TGTTTGCCCC CTTGAAGGAC TACTTTGCCT GTGAATAGTT

2851 CTACCCAGGA CTGGGGAGCT CTCGGTCAGA GCCAGTGCCC AGAGTCACCC

2901 CTCCCTGAAG GCACCTCACC TGTCCCCTTT TTAGCTCACC TGTGTGGGGC

2951 CTCACATCAC TGTACCTCAG CTTTCTCCTG CTCAGTGGGA GCAGAGCCTC

3001 CTGGCCCTTG CAGGGGAGCC CCGGTGCTCC CTCCGTGTGC ACAGCTCAGA

3051 CCTGGCTGCT TAGAGTAGCC CGGCATGGTG CTCATGTTCT CTTACCCTGA

3101 AACTTTAAAA CTTGAAGTAG GTAGTAAGAT GGCTTTCTTT TACCCTCCTG

3151 AGTTTATCAC TCAGAAGTGA TGGCTAAGAT ACCAAAAAAA CAAACAAAAA

3201 CAGAAACAAA AAACAAAAAA AAACCTCAAC AGCTCTCTTA GTACTCAGGT

3251 TCATGCTGCA AAATCACTTG AGATTTTGTT TTTAAGTAAC CCGTGCTCCA

3301 CATTTGCTGG AGGATGCTAT TGTGAATGTG GGCTCAGATG AGCAAGGTCA

3351 AGGGGCCAAA AAAAATTCCC CCTCTCCCCC CAGGAGTATT TGAAGATGAT

3401 GTTTATGGTT TAAGTCTTCC TGGCACCTTC CCCTTGCTTT GGTACAAGGG

FIG. 1B-3

3451 CTGAAGTCCT GTTGGTCTTG TAGCATTTCC CAGGATGATG ATGTCAGCAG

3501 GGATGACATC ACCACCTTTA GGGCTTTTCC CTGGCAGGGG CCCATGTGGC

3551 TAGTCCTCAC GAAGACTGGA GTAGAATGTT TGGAGCTCAG GAAGGGTGGG

3601 TGGAGTGGCC CTCTTCCAGG TGTGAGGGAT ACGAAGGAGG AAGCTTAGGG

3651 AAATCCATTC CCCACTCCCT CTTGCCAAAT GAGGGGCCCA GTCCCCAACA

3701 GCTCAGGTCC CCAGAACCCC CTAGTTCCTC ATGAGAAGCT AGGACCAGAA

3751 GCACATCGTT CCCCTTATCT GAGCAGTGTT TGGGGAACTA CAGTGAAAAC

3801 CTTCTGGAGA TGTTAAAAGC TTTTTACCCC ACGATAGATT GTGTTTTTAA

3851 GGGGTGCTTT TTTTAGGGGC ATCACTGGAG ATAAGAAAGC TGCATTTCAG

3901 AAATGCCATC GTAATGGTTT TTAAACACCT TTTACCTAAT TACAGGTGCT

3951 ATTTTATAGA AGCAGACAAC ACTTCTTTTT ATGACTCTCA GACTTCTATT

4001 TTCATGTTAC CATTTTTTTT GTAACTCGCA AGGTGTGGGC TTTTGTAACT

4051 TCACAGGTGT GGGGAGAGAC TGCCTTGTTT CAACAGTTTG TCTCCACTGG

4101 TTTCTAATTT TTAGGTGCAA AGATGACAGA TGCCCAGAGT TTACCTTTCT

4151 GGTTGATTAA AGTTGTATTT CTCTAAAAAA AAAAAAAAAA AAAAA

FIG. 1B-4

Human DNMT3A DNA Sequence

```
   1                                  GCCGCGG CACCAGGGCG CGCAGCCGGG
  28 CCGGCCCGAC CCCACCGGCC ATACGGTGGA GCCATCGAAG CCCCCACCCA
  78 CAGGCTGACA GAGGCACCGT TCACCAGAGG GCTCAACACC GGGATCTATG
 128 TTTAAGTTTT AACTCTCGCC TCCAAAGACC ACGATAATTC CTTCCCCAAA
 178 GCCCAGCAGC CCCCCAGCCC CGCGCAGCCC CAGCCTGCCT CCCGGCGCCC
 228 AGATGCCCGC CATGCCCTCC AGCGGCCCCG GGGACACCAG CAGCTCTGCT
 278 GCGGAGCCGG AGGAGGACCG AAAGGACGGA GAGGAGCAGG AGGAGCCGCG
 328 TGGCAAGGAG GAGCGCCAAG AGCCCAGCAC CACGGCACGG AAGGTGGGGC
 378 GGCCTGGGAG GAAGCGCAAG CACCCCCCGG TGGAAAGCGG TGACACGCCA
 428 AAGGACCCTG CGGTGATCTC CAAGTCCCCA TCCATGGCCC AGGACTCAGG
 478 CGCCTCAGAG CTATTACCCA ATGGGGACTT GGAGAAGCGG AGTGAGCCCC
 528 AGCCAGAGGA GGGGAGCCCT GCTGGGGGGC AGAAGGGCGG GGCCCCAGCA
 578 GAGGGAGAGG GTGCAGCTGA GACCCTGCCT GAAGCCTCAA GAGCAGTGGA
 628 AAATGGCTGC TGCACCCCCA AGGAGGGCCG AGGAGCCCCT GCAGAAGCGG
 678 GCAAAGAACA GAAGGAGACC AACATCGAAT CCATGAAAAT GGAGGGCTCC
 728 CGGGGCCGGC TGCGGGGTGG CTTGGGCTGG GAGTCCAGCC TCCGTCAGCG
 778 GCCCATGCCG AGGCTCACCT TCCAGGCGGG GGACCCCTAC TACATCAGCA
 828 AGCGCAAGCG GGACGAGTGG CTGGCACGCT GGAAAAGGGA GGCTGAGAAG
 878 AAAGCCAAGG TCAGTGCAGG AATGAATGCT GTGGAAGAAA ACCAGGGGCC
 928 CGGGGAGTCT CAGAAGGTGG AGGAGGCCAG CCCTCCTGCT GTGCAGCAGC
 978 CCACTGACCC CGCATCCCCC ACTGTGGCTA CCACGCCTGA GCCCGTGGGG
1028 TCCGATGCTG GGGACAAGAA TGCCACCAAA GCAGGCGATG ACGAGCCAGA
```

FIG. 1C-1

1078 GTACGAGGAC GGCCGGGGCT TTGGCATTGG GGAGCTGGTG TGGGGGAAAC

1128 TGCGGGGCTT CTCCTGGTGG CCAGGCCGCA TTGTGTCTTG GTGGATGACG

1178 GGCCGGAGCC GAGCAGCTGA AGGCACCCGC TGGGTCATGT GGTTCGGAGA

1228 CGGCAAATTC TCAGTGGTGT GTGTTGAGAA GCTGATGCCG CTGAGCTCGT

1278 TTTGCAGTGC GTTCCACCAG GCCACGTACA ACAAGCAGCC CATGTACCGC

1328 AAAGCCATCT ACGAGGTCCT GCAGGTGGCC AGCAGCCGCG CGGGGAAGCT

1378 GTTCCCGGTG TGCCACGACA GCGATGAGAG TGACACTGCC AAGGCCGTCG

1428 AGGTGCAGAA CAAGCCCATG ATTGAATGGG CCCTGGGGGG CTTCCAGCCT

1478 TCTGGCCCTA AGGGCCTGGA GCCACCAGAA GAAGAGAAGA ATCCCTACAA

1528 AGAAGTGTAC ACGGACATGT GGGTGGAACC TGAGGCAGCT GCCTACGCAC

1578 CACCTCCACC AGCCAAAAAG CCCCGGAAGA GCACAGCGGA GAAGCCCAAG

1628 GTCAAGGAGA TTATTGATGA GCGCACAAGA GAGCGGCTGG TGTACGAGGT

1678 GCGGCAGAAG TGCCGGAACA TTGAGGACAT CTGCATCTCC TGTGGGAGCC

1728 TCAATGTTAC CCTGGAACAC CCCCTCTTCG TTGGAGGAAT GTGCCAAAAC

1778 TGCAAGAACT GCTTTCTGGA GTGTGCGTAC CAGTACGACG ACGACGGCTA

1828 CCAGTCCTAC TGCACCATCT GCTGTGGGGG CCGTGAGGTG CTCATGTGCG

1878 GAAACAACAA CTGCTGCAGG TGCTTTTGCG TGGAGTGTGT GGACCTCTTG

1928 GTGGGGCCGG GGGCTGCCCA GGCAGCCATT AAGGAAGACC CCTGGAACTG

1978 CTACATGTGC GGGCACAAGG GTACCTACGG GCTGCTGCGG CGGCGAGAGG

2028 ACTGGCCCTC CCGGCTCCAG ATGTTCTTCG CTAATAACCA CGACCAGGAA

2078 TTTGACCCTC CAAAGGTTTA CCCACCTGTC CCAGCTGAGA AGAGGAAGCC

2128 CATCCGGGTG CTGTCTCTCT TTGATGGAAT CGCTACAGGG CTCCTGGTGC

2178 TGAAGGACTT GGGCATTCAG GTGGACCGCT ACATTGCCTC GGAGGTGTGT

FIG. 1C-2

2228 GAGGACTCCA TCACGGTGGG CATGGTGCGG CACCAGGGGA AGATCATGTA

2278 CGTCGGGGAC GTCCGCAGCG TCACACAGAA GCATATCCAG GAGTGGGGCC

2328 CATTCGATCT GGTGATTGGG GGCAGTCCCT GCAATGACCT CTCCATCGTC

2378 AACCCTGCTC GCAAGGGCCT CTACGAGGGC ACTGGCCGGC TCTTCTTTGA

2428 GTTCTACCGC CTCCTGCATG ATGCGCGGCC CAAGGAGGGA GATGATCGCC

2478 CCTTCTTCTG GCTCTTTGAG AATGTGGTGG CCATGGGCGT TAGTGACAAG

2528 AGGGACATCT CGCGATTTCT CGAGTCCAAC CCTGTGATGA TTGATGCCAA

2578 AGAAGTGTCA GCTGCACACA GGGCCCGCTA CTTCTGGGGT AACCTTCCCG

2628 GTATGAACAG GCCGTTGGCA TCCACTGTGA ATGATAAGCT GGAGCTGCAG

2678 GAGTGTCTGG AGCATGGCAG GATAGCCAAG TTCAGCAAAG TGAGGACCAT

2728 TACTACGAGG TCAAACTCCA TAAAGCAGGG CAAAGACCAG CATTTTCCTG

2778 TCTTCATGAA TGAGAAAGAG GACATCTTAT GGTGCACTGA AATGGAAAGG

2828 GTATTTGGTT TCCCAGTCCA CTATACTGAC GTCTCCAACA TGAGCCGCTT

2878 GGCGAGGCAG AGACTGCTGG GCCGGTCATG GAGCGTGCCA GTCATCCGCC

2928 ACCTCTTCGC TCCGCTGAAG GAGTATTTTG CGTGTGTGTA AGGGACATGG

2978 GGGCAAACTG AGGTAGCGAC ACAAAGTTAA ACAAACAAAC AAAAAACACA

3028 AAACATAATA AAACACCAAG AACATGAGGA TGGAGAGAAG TATCAGCACC

3078 CAGAAGAGAA AAAGGAATTT AAAACAAAAA CCACAGAGGC GGAAATACCG

3128 GAGGGCTTTG CCTTGCGAAA AGGGTTGGAC ATCATCTCCT GATTTTTCAA

3178 TGTTATTCTT CAGTCCTATT TAAAAACAAA ACCAAGCTCC CTTCCCTTCC

3228 TCCCCCTTCC CTTTTTTTTC GGTCAGACCT TTTATTTTCT ACTCTTTTCA

3278 GAGGGGTTTT CTGTTTGTTT GGGTTTTGTT TCTTGCTGTG ACTGAAACAA

3328 GAAGGTTATT GCAGCAAAAA TCAGTAACAA AAAATAGTAA CAATACCTTG

3378 CAGAGGAAAG GTGGGAGGAG AGGAAAAAAG GGAAATTTTT AAAGAAATCT

FIG. 1C-3

3428 ATATATTGGG TTGTTTTTTT TTTTGTTTTT TGTTTTTTTT TTTTGGGTTT

3478 TTTTTTTTTA CTATATATCT TTTTTTTGTT GTCTCTAGCC TGATCAGATA

3528 GGAGCACAAG CAGGGGACGG AAAGAGAGAG ACACTCAGGC GGCAGCATTC

3578 CCTCCCAGCC ACTGAGCTGT CGTGCCAGCA CCATTCCTGG TCACGCAAAA

3628 CAGAACCCAG TTAGCAGCAG GGAGACGAGA ACACCACACA AGACATTTTT

3678 CTACAGTATT TCAGGTGCCT ACCACACAGG AAACCTTGAA GAAAATCAGT

3728 TTCTAGAAGC CGCTGTTACC TCTTGTTTAC AGTTTATATA TATATGATAG

3778 ATATGAGATA TATATATAAA AGGTACTGTT AACTACTGTA CAACCCGACT

3828 TCATAATGGT GCTTTCAAAC AGCGAGATGA GTAAAAACAT CAGCTTCCAC

3878 GTTGCCTTCT GCGCAAAGGG TTTCACCAAG GATGGAGAAA GGGAGACAGC

3928 TTGCAGATGG CGCGTTCTCA CGGTGGGCTC TTCCCCTTGG TTTGTAACGA

3978 AGTGAAGGAG GAGAACTTGG GAGCCAGGTT CTCCCTGCCA AAAAGGGGGC

4028 TAGATGAGGT GGTCGGGCCC GTGGACAGCT GAGAGTGGGA TTCATCCAGA

4078 CTCATGCAAT AACCCTTTGA TTGTTTTCTA AAAGGAGACT CCCTCGGCAA

4128 GATGGCAGAG GGTACGGAGT CTTCAGGCCC AGTTTCTCAC TTTAGCCAAT

4178 TCGAGGGCTC CTTGTGGTGG GATCAGAACT AATCCAGAGT GTGGGAAAGT

4228 GACAGTCAAA ACCCCACCTG GAGCAAATAA AAAAACATAC AAAACGTAAA

4278 AAAAAAAAAA AAAAAA

FIG. 1C-4

Human DNMT3B1 DNA Sequence:

```
   1  GGCCGCGAAT TCGGCACGAG CCCTGCACGG CCGCCAGCCG GCCTCCCGCC
  51  AGCCAGCCCC GACCCGCGGC TCCGCCGCCC AGCCGCGCCC CAGCCAGCCC
 101  TGCGGCAGGA AAGCATGAAG GGAGACACCA GGCATCTCAA TGGAGAGGAG
 151  GACGCCGGCG GGAGGGAAGA CTCGATCCTC GTCAACGGGG CCTGCAGCGA
 201  CCAGTCCTCC GACTCGCCCC CAATCCTGGA GGCTATCCGC ACCCCGGAGA
 251  TCAGAGGCCG AAGATCAAGC TCGCGACTCT CCAAGAGGGA GGTGTCCAGT
 301  CTGCTAAGCT ACACACAGGA CTTGACAGGC GATGGCGACG GGGAAGATGG
 351  GGATGGCTCT GACACCCCAG TCATGCCAAA GCTCTTCCGG GAAACCAGGA
 401  CTCGTTCAGA AAGCCCAGCT GTCCGAACTC GAAATAACAA CAGTGTCTCC
 451  AGCCGGGAGA GGCACAGGCC TTCCCCACGT TCCACCCGAG GCCGGCAGGG
 501  CCGCAACCAT GTGGACGAGT CCCCCGTGGA GTTCCCGGCT ACCAGGTCCC
 551  TGAGACGGCG GGCAACAGCA TCGGCAGGAA CGCCATGGCC GTCCCCTCCC
 601  AGCTCTTACC TTACCATCGA CCTCACAGAC GACACAGAGG ACACACATGG
 651  GACGCCCCAG AGCAGCAGTA CCCCCTACGC CCGCCTAGCC CAGGACAGCC
 701  AGCAGGGGGG CATGGAGTCC CCGCAGGTGG AGGCAGACAG TGGAGATGGA
 751  GACAGTTCAG AGTATCAGGA TGGGAAGGAG TTTGGAATAG GGGACCTCGT
 801  GTGGGGAAAG ATCAAGGGCT TCTCCTGGTG GCCCGCCATG GTGGTGTCTT
 851  GGAAGGCCAC CTCCAAGCGA CAGGCTATGT CTGGCATGCG GTGGGTCCAG
 901  TGGTTTGGCG ATGGCAAGTT CTCCGAGGTC TCTGCAGACA AACTGGTGGC
 951  ACTGGGGCTG TTCAGCCAGC ACTTTAATTT GGCCACCTTC AATAAGCTCG
1001  TCTCCTATCG AAAAGCCATG TACCATGCTC TGGAGAAAGC TAGGGTGCGA
1051  GCTGGCAAGA CCTTCCCCAG CAGCCCTGGA GACTCATTGG AGGACCAGCT
1101  GAAGCCCATG TTGGAGTGGG CCCACGGGGG CTTCAAGCCC ACTGGGATCG
1151  AGGGCCTCAA ACCCAACAAC ACGCAACCAG TGGTTAATAA GTCGAAGGTG
```

FIG. 1D-1

1201 CGTCGTGCAG GCAGTAGGAA ATTAGAATCA AGGAAATACG AGAACAAGAC

1251 TCGAAGACGC ACAGCTGACG ACTCAGCCAC CTCTGACTAC TGCCCCGCAC

1301 CCAAGCGCCT CAAGACAAAT TGCTATAACA ACGGCAAAGA CCGAGGGGAT

1351 GAAGATCAGA GCCGAGAACA AATGGCTTCA GATGTTGCCA ACAACAAGAG

1401 CAGCCTGGAA GATGGCTGTT TGTCTTGTGG CAGGAAAAAC CCCGTGTCCT

1451 TCCACCCTCT CTTTGAGGGG GGGCTCTGTC AGACATGCCG GGATCGCTTC

1501 CTTGAGCTGT TTTACATGTA TGATGACGAT GGCTATCAGT CTTACTGCAC

1551 TGTGTGCTGC GAGGGCCGAG AGCTGCTGCT TTGCAGCAAC ACGAGCTGCT

1601 GCCGGTGTTT CTGTGTGGAG TGCCTGGAGG TGCTGGTGGG CACAGGCACA

1651 GCGGCCGAGG CCAAGCTTCA GGAGCCCTGG AGCTGCTACA TGTGTCTCCC

1701 GCAGCGCTGT CATGGCGTCC TGCGGCGCCG GAAGGACTGG AACGTGCGCC

1751 TGCAGGCCTT CTTCACCAGT GACACGGGGC TTGAATACGA AGCCCCCAAG

1801 CTGTACCCTG CCATTCCCGC AGCCCGAAGG CGGCCCATTC GAGTCCTGTC

1851 ATTGTTTGAT GGCATCGCGA CAGGCTACCT AGTCCTCAAA GAGTTGGGCA

1901 TAAAGGTAGG AAAGTACGTC GCTTCTGAAG TGTGTGAGGA GTCCATTGCT

1951 GTTGGAACCG TGAAGCACGA GGGGAATATC AAATACGTGA ACGACGTGAG

2001 GAACATCACA AAGAAAAATA TTGAAGAATG GGGCCCATTT GACTTGGTGA

2051 TTGGCGGAAG CCCATGCAAC GATCTCTCAA ATGTGAATCC AGCCAGGAAA

2101 GGCCTGTATG AGGGTACAGG CCGGCTCTTC TTCGAATTTT ACCACCTGCT

2151 GAATTACTCA CGCCCCAAGG AGGGTGATGA CCGGCCGTTC TTCTGGATGT

2201 TTGAGAATGT TGTAGCCATG AAGGTTGGCG ACAAGAGGGA CATCTCACGG

2251 TTCCTGGAGT GTAATCCAGT GATGATTGAT GCCATCAAAG TTTCTGCTGC

2301 TCACAGGGCC CGATACTTCT GGGGCAACCT ACCCGGGATG AACAGGCCCG

2351 TGATAGCATC AAAGAATGAT AAACTCGAGC TGCAGGACTG CTTGGAATAC

2401 AATAGGATAG CCAAGTTAAA GAAAGTACAG ACAATAACCA CCAAGTCGAA

FIG. 1D-2

2451 CTCGATCAAA CAGGGGAAAA ACCAACTTTT CCCTGTTGTC ATGAATGGCA

2501 AAGAAGATGT TTTGTGGTGC ACTGAGCTCG AAAGGATCTT TGGCTTTCCT

2551 GTGCACTACA CAGACGTGTC CAACATGGGC CGTGGTGCCC GCCAGAAGCT

2601 GCTGGGAAGG TCCTGGAGCG TGCCTGTCAT CCGACACCTC TTCGCCCCTC

2651 TGAAGGACTA CTTTGCATGT GAATAGTTCC AGCCAGGCCC CAAGCCCACT

2701 GGGGTGTGTG GCAGAGCCAG GACCCAGGAG GTGTGATTCC TGAAGGCATC

2751 CCCAGGCCCT GCTCTTCCTC AGCTGTGTGG GTCATACCGT GTACCTCAGT

2801 TCCCTCTTGC TCAGTGGGGG CAGAGCCACC TGACTCTTGC AGGGGTAGCC

2851 TGAGGTGCCG CCTCCTTGTG CACAAATCAG ACCTGGCTGC TTGGAGCAGC

2901 CTAACACGGT GCTCATTTTT TCTTCTCCTA AAACTTTAAA ACTTGAAGTA

2951 GGTAGCAACG TGGCTTTTTT TTTTTCCCTT CCTGGGTCTA CCACTCAGAG

3001 AAACAATGGC TAAGATACCA AAACCACAGT GCCGACAGCT CTCCAATACT

3051 CAGGTTAATG CTGAAAAATC ATCCAAGACA GTTATTGCAA GAGTTTAATT

3101 TTTGAAAACT GGGTACTGCT ATGTGTTTAC AGACGTGTGC AGTTGTAGGC

3151 ATGTAGCTAC AGGACATTTT TAAGGGCCCA GGATCGTTTT TTCCCAGGGC

3201 AAGCAGAAGA GAAAATGTTG TATATGTCTT TTACCCGGCA CATTCCCCTT

3251 GCCTAAATAC AAGGGCTGGA GTCTGCACGG GACCTATTAG AGTATTTTCC

3301 ACAATGATGA TGATTTCAGC AGGGATGACG TCATCATCAC ATTCAGGGCT

3351 ATTTTTTCCC CCACAAACCC AAGGGCAGGG GCCACTCTTA GCTAAATCCC

3401 TCCCCGTGAC TGCAATAGAA CCCTCTGGGG AGCTCAGGAA GGGGTGTGCT

3451 GAGTTCTATA ATATAAGCTG CCATATATTT TGTAGACAAG TATGGCTCCT

3501 CCATATCTCC CTCTTCCCTA GGAGAGGAGT GTGAAGCAAG GAGCTTAGAT

3551 AAGACACCCC CTCAAACCCA TTCCCTCTCC AGGAGACCTA CCCTCCACAG

3601 GCACAGGTCC CCAGATGAGA AGTCTGCTAC CCTCATTTCT CATCTTTTTA

3651 CTAAACTCAG AGGCAGTGAC AGCAGTCAGG GACAGACATA CATTTCTCAT

FIG. 1D-3

3701 ACCTTCCCCA CATCTGAGAG ATGACAGGGA AAACTGCAAA GCTCGGTGCT

3751 CCCTTTGGAG ATTTTTTAAT CCTTTTTTAT TCCATAAGAA GTCGTTTTTA

3801 GGGAGAACGG GAATTCAGAC AAGCTGCATT TCAGAAATGC TGTCATAATG

3851 GTTTTTAACA CCTTTTACTC TTCTTACTGG TGCTATTTTG TAGAATAAGG

3901 AACAACGTTG ACAAGTTTTG TGGGGCTTTT TATACACTTT TTAAAATCTC

3951 AAACTTCTAT TTTTATGTTT AACGTTTTCA TTAAAATTTT TTTGTAACTG

4001 GAGCCACGAC GTAACAAATA TGGGGAAAAA ACTGTGCCTT GTTTCAACAG

4051 TTTTTGCTAA TTTTTAGGCT GAAAGATGAC GGATGCCTAG AGTTTACCTT

4101 ATGTTTAATT AAAATCAGTA TTTGTCTAAA AAAAAAAAAA AAAAA

FIG. 1D-4

Mouse Dnmt3a Protein

1 MPSSGPGDTS SSSLEREDDR KEGEEQEENR GKEERQEPSA TARKVGRPGR

51 KRKHPPVESS DTPKDPAVTT KSQPMAQDSG PSDLLPNGDL EKRSEPQPEE

101 GSPAAGQKGG APAEGEGTET PPEASRAVEN GCCVTKEGRG ASAGEGKEQK

151 QTNIESMKME GSRGRLRGGL GWESSLRQRP MPRLTFQAGD PYYISKRKRD

201 EWLARWKREA EKKAKVIAVM NAVEENQASG ESQKVEEASP PAVQQPTDPA

251 SPTVATTPEP VGGDAGDKNA TKAADDEPEY EDGRGFGIGE LVWGKLRGFS

301 WWPGRIVSWW MTGRSRAAEG TRWVMWFGDG KFSVVCVEKL MPLSSFCSAF

351 HQATYNKQPM YRKAIYEVLQ VASSRAGKLF PACHDSDESD SGKAVEVQNK

401 QMIEWALGGF QPSGPKGLEP PEEEKNPYKE VYTDMWVEPE AAAYAPPPPA

451 KKPRKSTTEK PKVKEIIDER TRERLVYEVR QKCRNIEDIC ISCGSLNVTL

501 EHPLFIGGMC QNCKNCFLEC AYQYDDDGYQ SYCTICCGGR EVLMCGNNNC

551 CRCFCVECVD LLVGPGAAQA AIKEDPWNCY MCGHKGTYGL LRRREDWPSR

601 LQMFFANNHD QEFDPPKVYP PVPAEKRKPI RVLSLFDGIA TGLLVLKDLG

651 IQVDRYIASE VCEDSITVGM VRHQGKIMYV GDVRSVTQKH IQEWGPFDLV

701 IGGSPCNDLS IVNPARKGLY EGTGRLFFEF YRLLHDARPK EGDDRPFFWL

751 FENVVAMGVS DKRDISRFLE SNPVMIDAKE VSAAHRARYF WGNLPGMNRP

801 LASTVNDKLE LQECLEHGRI AKFSKVRTIT TRSNSIKQGK DQHFPVFMNE

851 KEDILWCTEM ERVFGFPVHY TDVSNMSRLA RQRLLGRSWS VPVIRHLFAP

901 LKEYFACV*

FIG. 2A

Mouse Dnmt3b1 Protein

1 MKGDSRHLNE EEGASGYEEC IIVNGNFSDQ SSDTKDAPSP PVLEAICTEP

51 VCTPETRGRR SSSRLSKREV SSLLNYTQDM TGDGDRDDEV DDGNGSDILM

101 PKLTRETKDT RTRSESPAVR TRHSNGTSSL ERQRASPRIT RGRQGRHHVQ

151 EYPVEFPATR SRRRRASSSA STPWSSPASV DFMEEVTPKS VSTPSVDLSQ

201 DGDQEGMDTT QVDAESRDGD STEYQDDKEF GIGDLVWGKI KGFSWWPAMV

251 VSWKATSKRQ AMPGMRWVQW FGDGKFSEIS ADKLVALGLF SQHFNLATFN

301 KLVSYRKAMY HTLEKARVRA GKTFSSSPGE SLEDQLKPML EWAHGGFKPT

351 GIEGLKPNKK QPVVNKSKVR RSDSRNLEPR RRENKSRRRT TNDSAASESP

401 PPKRLKTNSY GGKDRGEDEE SRERMASEVT NNKGNLEDRC LSCGKKNPVS

451 FHPLFEGGLC QSCRDRFLEL FYMYDEDGYQ SYCTVCCEGR ELLLCSNTSC

501 CRCFCVECLE VLVGAGTAED AKLQEPWSCY MCLPQRCHGV LRRRKDWNMR

551 LQDFFTTDPD LEEFEPPKLY PAIPAAKRRP IRVLSLFDGI ATGYLVLKEL

601 GIKVEKYIAS EVCAESIAVG TVKHEGQIKY VNDVRKITKK NIEEWGPFDL

651 VIGGSPCNDL SNVNPARKGL YEGTGRLFFE FYHLLNYTRP KEGDNRPFFW

701 MFENVVAMKV NDKKDISRFL ACNPVMIDAI KVSAAHRARY FWGNLPGMNR

751 PVMASKNDKL ELQDCLEFSR TAKLKKVQTI TTKSNSIRQG KNQLFPVVMN

801 GKDDVLWCTE LERIFGFPAH YTDVSNMGRG ARQKLLGRSW SVPVIRHLFA

851 PLKDYFACE*

FIG. 2B

Human DNMT3A Protein

1 MPAMPSSGPG DTSSSAAERE EDRKDGEEQE EPRGKEERQE PSTTARKVGR

51 PGRKRKHPPV ESGDTPKDPA VISKSPSMAQ DSGASELLPN GDLEKRSEPQ

101 PEEGSPAGGQ KGGAPAEGEG AAETLPEASR AVENGCCTPK EGRGAPAEAG

151 KEQKETNIES MKMEGSRGRL RGGLGWESSL RQRPMPRLTF QAGDPYYISK

201 RKRDEWLARW KREAEKKAKV IAGMNAVEEN QGPGESQKVE EASPPAVQQP

251 TDPASPTVAT TPEPVGSDAG DKNATKAGDD EPEYEDGRGF GIGELVWGKL

301 RGFSWWPGRI VSWWMTGRSR AAEGTRWVMW FGDGKFSVVC VEKLMPLSSF

351 CSAFHQATYN KQPMYRKAIY EVLQVASSRA GKLFPVCHDS DESDTAKAVE

401 VQNKPMIEWA LGGFQPSGPK GLEPPEEEKN PYKEVYTDMW VEPEAAAYAP

451 PPPAKKPRKS TAEKPKVKEI IDERTRERLV YEVRQKCRNI EDICISCGSL

501 NVTLEHPLFV GGMCQNCKNC FLECAYQYDD DGYQSYCTIC CGGREVLMCG

551 NNNCCRCFCV ECVDLLVGPG AAQAAIKEDP WNCYMCGHKG TYGLLRRRED

601 WPSRLQMFFA NNHDQEFDPP KVYPPVPAEK RKPIRVLSLF DGIATGLLVL

651 KDLGIQVDRY IASEVCEDSI TVGMVRHQGK IMYVGDVRSV TQKHIQEWGP

701 FDLVIGGSPC NDLSIVNPAR KGLYEGTGRL FFEFYRLLHD ARPKEGDDRP

751 FFWLFENVVA MGVSDKRDIS RFLESNPVMI DAKEVSAAHR ARYFWGNLPG

801 MNRPLASTVN DKLELQECLE HGRIAKFSKV RTITTRSNSI KQGKDQHFPV

851 FMNEKEDILW CTEMERVFGF PVHYTDVSNM SRLARQRLLG RSWSVPVIRH

901 LFAPLKEYFA CV*

FIG. 2C

Human DNMT3B1 Protein

```
  1 MKGDTRHLNG EEDAGGREDS ILVNGACSDQ SSDSPPILEA IRTPEIRGRR
 51 SSSRLSKREV SSLLSYTQDL TGDGDGEDGD GSDTPVMPKL FRETRTRSES
101 PAVRTRNNNS VSSRERHRPS PRSTRGRQGR NHVDESPVEF PATRSLRRRA
151 TASAGTPWPS PPSSYLTIDL TDDTEDTHGT PQSSSTPYAR LAQDSQQGGM
201 ESPQVEADSG DGDSSEYQDG KEFGIGDLVW GKIKGFSWWP AMVVSWKATS
251 KRQAMSGMRW VQWFGDGKFS EVSADKLVAL GLFSQHFNLA TFNKLVSYRK
301 AMYHALEKAR VRAGKTFPSS PGDSLEDQLK PMLEWAHGGF KPTGIEGLKP
351 NNTQPVVNKS KVRRAGSRKL ESRKYENKTR RRTADDSATS DYCPAPKRLK
401 TNCYNNGKDR GDEDQSREQM ASDVANNKSS LEDGCLSCGR KNPVSFHPLF
451 EGGLCQTCRD RFLELFYMYD DDGYQSYCTV CCEGRELLLC SNTSCCRCFC
501 VECLEVLVGT GTAAEAKLQE PWSCYMCLPQ RCHGVLRRRK DWNVRLQAFF
551 TSDTGLEYEA PKLYPAIPAA RRRPIRVLSL FDGIATGYLV LKELGIKVGK
601 YVASEVCEES IAVGTVKHEG NIKYVNDVRN ITKKNIEEWG PFDLVIGGSP
651 CNDLSNVNPA RKGLYEGTGR LFFEFYHLLN YSRPKEGDDR PFFWMFENVV
701 AMKVGDKRDI SRFLECNPVM IDAIKVSAAH RARYFWGNLP GMNRPVIASK
751 NDKLELQDCL EYNRIAKLKK VQTITTKSNS IKQGKNQLFP VVMNGKEDVL
801 WCTELERIFG FPVHYTDVSN MGRGARQKLL GRSWSVPVIR HLFAPLKDYF
851 ACE*
```

FIG. 2D

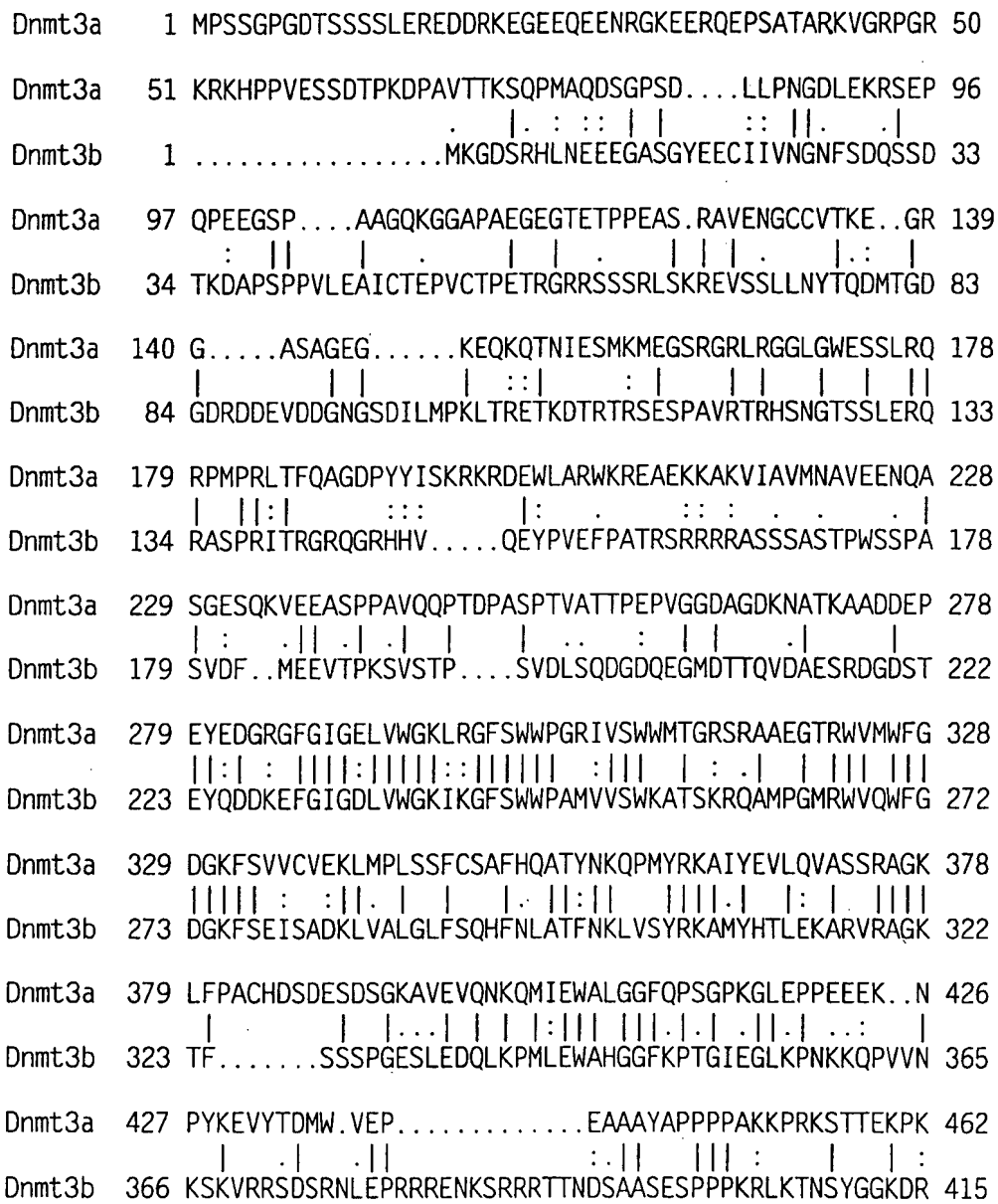

Dnmt3a 1 MPSSGPGDTSSSSLEREDDRKEGEEQEENRGKEERQEPSATARKVGRPGR 50
Dnmt3a 51 KRKHPPVESSDTPKDPAVTTKSQPMAQDSGPSD....LLPNGDLEKRSEP 96
Dnmt3b 1 .....MKGDSRHLNEEEGASGYEECIIVNGNFSDQSSD 33
Dnmt3a 97 QPEEGSP....AAGQKGGAPAEGEGTETPPEAS.RAVENGCCVTKE..GR 139
Dnmt3b 34 TKDAPSPPVLEAICTEPVCTPETRGRRSSSRLSKREVSSLLNYTQDMTGD 83
Dnmt3a 140 G.....ASAGEG.....KEQKQTNIESMKMEGSRGRLRGGLGWESSLRQ 178
Dnmt3b 84 GDRDDEVDDGNGSDILMPKLTRETKDTRTRSESPAVRTRHSNGTSSLERQ 133
Dnmt3a 179 RPMPRLTFQAGDPYYISKRKRDEWLARWKREAEKKAKVIAVMNAVEENQA 228
Dnmt3b 134 RASPRITRGRQGRHHV.....QEYPVEFPATRSRRRRASSSASTPWSSPA 178
Dnmt3a 229 SGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP 278
Dnmt3b 179 SVDF..MEEVTPKSVSTP....SVDLSQDGDQEGMDTTQVDAESRDGDST 222
Dnmt3a 279 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFG 328
Dnmt3b 223 EYQDDKEFGIGDLVWGKIKGFSWWPAMVVSWKATSKRQAMPGMRWVQWFG 272
Dnmt3a 329 DGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGK 378
Dnmt3b 273 DGKFSEISADKLVALGLFSQHFNLATFNKLVSYRKAMYHTLEKARVRAGK 322
Dnmt3a 379 LFPACHDSDESDSGKAVEVQNKQMIEWALGGFQPSGPKGLEPPEEEK..N 426
Dnmt3b 323 TF.....SSSPGESLEDQLKPMLEWAHGGFKPTGIEGLKPNKKQPVVN 365
Dnmt3a 427 PYKEVYTDMW.VEP.....EAAAYAPPPPAKKPRKSTTEKPK 462
Dnmt3b 366 KSKVRRSDSRNLEPRRRENKSRRRTTNDSAASESPPPKRLKTNSYGGKDR 415

FIG.3A-1


Dnmt3a 463 VKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPFFIGGMCQN 512
Dnmt3b 416 GE...DEESRERMASEVTNNKGNLEDRCLSCGKKNPVSFHPLFEGGLCQS 462
Dnmt3a 513 CKNCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLL 562
Dnmt3b 463 CRDRFLELFYMYDEDGYQSYCTVCCEGRELLLCSNTSCCRCFCVECLEVL 512
Dnmt3a 563 VGPGAAQAAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHD.Q 611
Dnmt3b 513 VGAGTAEDAKLQEPWSCYMCLPQRCHGVLRRRKDWNMRLQDFFTTDPDLE 562
Dnmt3a 612 EFDPPKVYPPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEV 661
Dnmt3b 563 EFEPPKLYPAIPAAKRRPIRVLSLFDGIATGYLVLKELGIKVEKYIASEV 612
Dnmt3a 662 CEDSITVGMVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSI 711
Dnmt3b 613 CAESIAVGTVKHEGQIKYVNDVRKITKKNIEEWGPFDLVIGGSPCNDLSN 662
Dnmt3a 712 VNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSD 761
Dnmt3b 663 VNPARKGLYEGTGRLFFEFYHLLNYTRPKEGDNRPFFWMFENVVAMKVND 712
Dnmt3a 762 KRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLEL 811
Dnmt3b 713 KKDISRFLACNPVMIDAIKVSAAHRARYFWGNLPGMNRPVMASKNDKLEL 762
Dnmt3a 812 QECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEME 861
Dnmt3b 763 QDCLEFSRTAKLKKVQTITTKSNSIRQGKNQLFPVVMNGKDDVLWCTELE 812
Dnmt3a 862 RVFGFPVHYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV* 909
Dnmt3b 813 RIFGFPAHYTDVSNMGRGARQKLLGRSWSVPVIRHLFAPLKDYFACE* 860

FIG.3A-2


DNMT3A 1 MPAMPSSGPGDTSSSAAEREEDRKDGEEQEEPRGKEERQEPSTTARKVGR
DNMT3A 51 PGRKRKHPPVESGDTPKDPAVISKSPSMAQDSGASELLPNGDLEKRSEPQ
DNMT3B 1 .....MKGDTRHLNGEEDAGGREDSILVNGACSDQSSDSP
DNMT3A 101 PEEGSPAGGQKGGAPAEGEGAAETLPEASRAVENGCCTPKEGRGAPAEAG
DNMT3B 36 PILEAIRTPEIRGGWASSRLSKREVSSLLSYTQDLTGDGDGEDGDGSDTP
DNMT3A 151 KEQKETNIESMKMEGSRGRLRGGLGWESSLRQRPMPRLTFQAGDPYYISK
DNMT3B 86 VMPKLFRETRTRSESPAVRTRNNNSVSSRERHRPSPRSTRGRQGRNHVDE
DNMT3A 201 RKRDEWLARWKREAEKKAKVIAGMNAVEENQGPGESQKVEEASPPAVQQP
DNMT3B 136 SPVEFPATRSLRRRATASAGTPWPSPPSSYLTIDLTDDTEDTH..GTPQS
DNMT3A 251 TDPASPTVATTPEPVGSDAGDKNATKAGDDEPEYEDGRGFGIGELVWGKL
DNMT3B 184 SSTPYARLAQDSQQGGMESPQVEADSGDGDSSEYQDGKEFGIGDLVWGKI
DNMT3A 301 RGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCVEKLMPLSSF
DNMT3B 234 KGFSWWPAMVVSWKATSKRQAMSGMRWVQWFGDGKFSEVSADKLVALGLF
DNMT3A 351 CSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE
DNMT3B 284 SQHFNLATFNKLVSYRKAMYHALEKARVRAGKTFP.....SSPGDSLE
DNMT3A 401 VQNKPMIEWALGGFQPSGPKGLEP....PEEEKNPYKEVYTDMWVE....
DNMT3B 327 DQLKPMLEWAHGGFKPTGIEGLKPNNTQPVVNKSKVRRAGSRKLESRKYE
DNMT3A 443 .....PEAAAYAPPPPAKKPRKSTAEKPKVKEIIDERTRERLVYEVRQ
DNMT3B 377 NKTRRRTADDSATSDYCPAPKRLKTNCYNNGKDRGDEDQSREQMASDVAN

FIG.3B-1


DNMT3A 486 KCRNIEDICISCGSLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQS
DNMT3B 427 NKSSLEDGCLSCGRKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDGYQS
DNMT3A 536 YCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKEDPWNCYM
DNMT3B 477 YCTVCCEGRELLLCSNTSCCRCFCVECLEVLVGTGTAAEAKLQEPWSCYM
DNMT3A 586 CGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIR
DNMT3B 527 CLPQRCHGVLRRRKDWNVRLQAFFTSDTGLEYEAPKLYPAIPAARRRPIR
DNMT3A 636 VLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVG
DNMT3B 577 VLSLFDGIATGYLVLKELGIKVGKYVASEVCEESIAVGTVKHEGNIKYVN
DNMT3A 686 DVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFY
DNMT3B 627 DVRNITKKNIEEWGPFDLVIGGSPCNDLSNVNPARKGLYEGTGRLFFEFY
DNMT3A 736 RLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFLESNPVMIDAKEV
DNMT3B 677 HLLNYSRPKEGDDRPFFWMFENVVAMKVGDKRDISRFLECNPVMIDAIKV
DNMT3A 786 SAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGRIAKFSKVRTITT
DNMT3B 727 SAAHRARYFWGNLPGMNRPVIASKNDKLELQDCLEYNRIAKLKKVQTITT
DNMT3A 836 RSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVHYTDVSNMSRLAR
DNMT3B 777 KSNSIKQGKNQLFPVVMNGKEDVLWCTELERIFGFPVHYTDVSNMGRGAR
DNMT3A 886 QRLLGRSWSVPVIRHLFAPLKEYFACV*
DNMT3B 827 QKLLGRSWSVPVIRHLFAPLKDYFACE*

FIG.3B-2

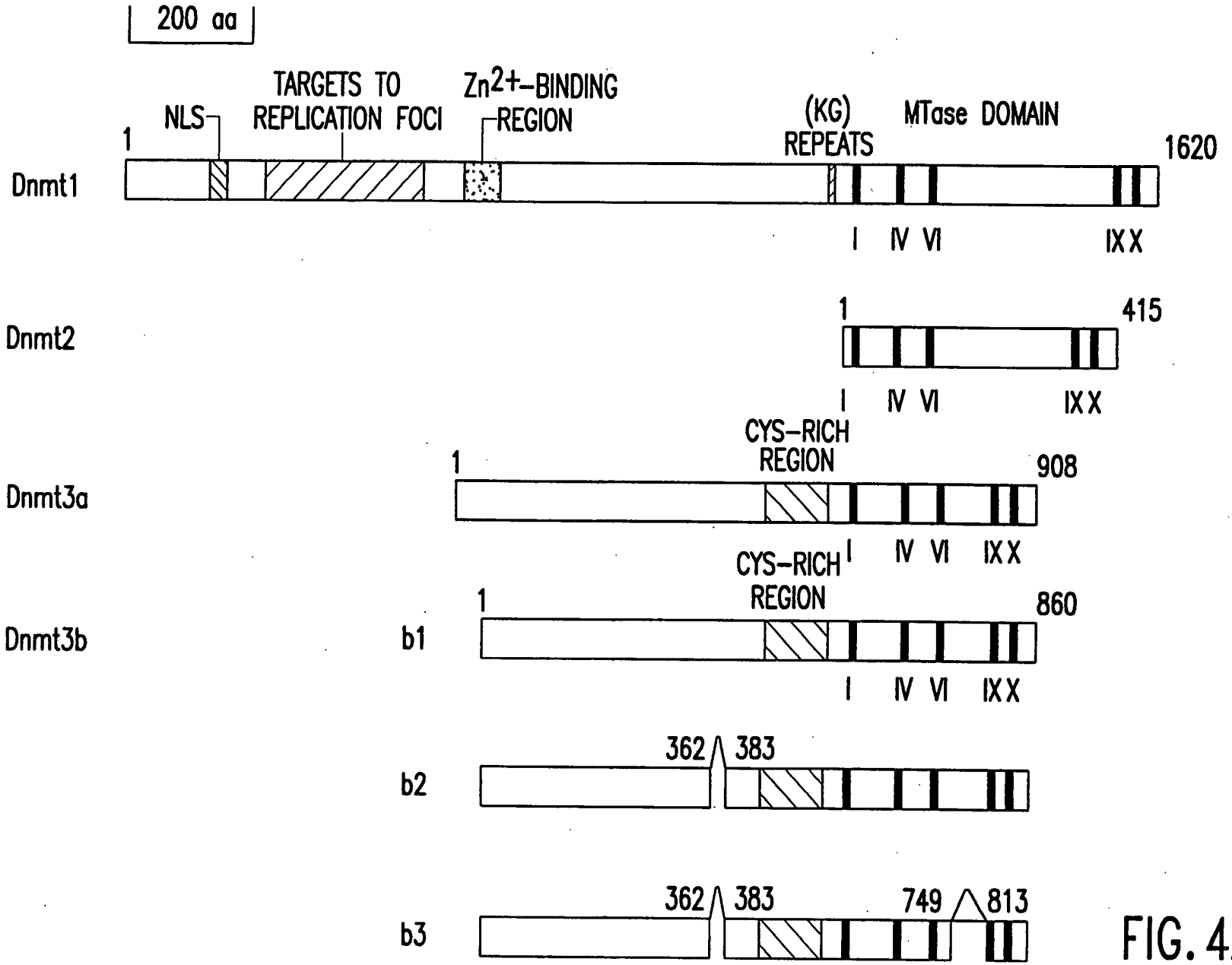
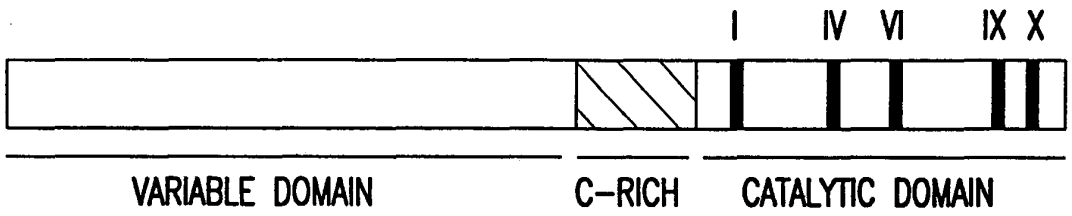

200 aa
TARGETS TO REPLICATION FOCI
Zn2+-BINDING REGION
NLS
(KG) REPEATS
MTase DOMAIN
1
1620
Dnmt1
I IV VI IXX
1
415
Dnmt2
I IV VI IXX
CYS-RICH REGION
1
908
Dnmt3a
I IV VI IXX
CYS-RICH REGION
1
860
Dnmt3b
b1
I IV VI IXX
362
383
b2
362
383
749
813
b3

FIG. 4A



DNMT3A	100%	100%	100%
DNMT3B	28%	57%	81%
Zmt3	23%	52%	70%

FIG. 4B

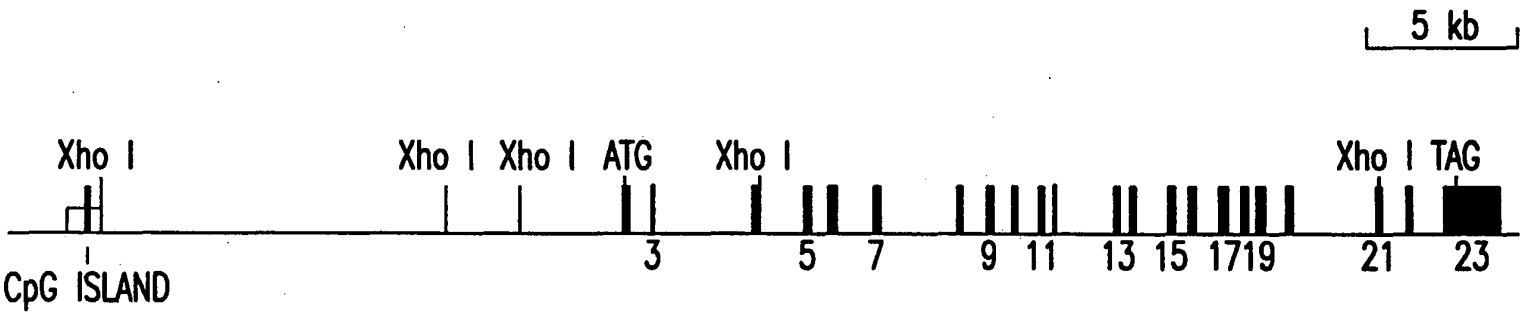

5 kb
Xho I
Xho I
Xho I
ATG
Xho I
Xho I
TAG
CpG ISLAND
3
5
7
9
11
13
15
17
19
21
23

FIG. 4C

```
Exon1  (>=90bp)CGGCAGgtgagcgccccgggg.intron(17618bp).tggcttctcccacagGAAAGC
Exon2  ( 148bp)TCAGAGgtggctgggcagtgg.intron(  887bp).CTGTTTCCTCTACAGGCCGAA
Exon3  (  62bp)ACACAGgtatggtctctgctc.intron( 3343bp).tgtttccttataaagGACTTG
Exon4  ( 102bp)CCAGCTgtaagtagccacacc.intron( 1642bp).ctctcttgcttctagGTCCGA
Exon5  ( 125bp)ACCAGGgttgttccccagatg.intron(  602bp).tccttctgtccacagTCCCTG
Exon6  ( 222bp)TATCAGgtatggccgagaggg.intron( 1403bp).tgggttttcttccagGATGGG
Exon7  ( 159bp)TCCGAGgtgagtccggggaag.intron( 2588bp).gtctttctcttttagGTCTCT
Exon8  ( 108bp)CTGGAGgtaacatgggatgag.intron(  917bp).actctgcctttgcagAAAGCT
Exon9  ( 145bp)AACCAGgtgggaatgagtccc.intron(  765bp).ttttcccctcaaaagTGGTTA
Exon10(  60bp)AATACGgtatttccttcctgt.intron( 1813bp).aattacctttcacagAGAACA
Exon11( 126bp)GCCGAGgtgattgttgggtac.intron(  115bp).ttcttttctcaatagAACAAA
Exon12(  45bp)TGGAAGgtaacgttctctccc.intron( 1095bp).ctgtttttcttacagATGGCT
Exon13(  80bp)TGCCGGgtaagtcctcctact.intron(  417bp).ctctctggctgccagGATCGC
Exon14( 113bp)CTGCCGgtgagcactgggccc.intron( 1160bp).tgccactgggtccagGTGTTT
Exon15( 184bp)GAATACgtaagccacaggctc.intron(  600bp).ttccttacctggcagGAAGCC
Exon16(  85bp)CGACAGgtgagttcggggaac.intron(  824bp).ctctggcccccacagGCTACC
Exon17( 146bp)AAAAATgtgagggcagtctgt.intron(  536bp).gtctctctctttcagATTGAA
Exon18(  91bp)TGTATGgtgagcatccttctc.intron(  352bp).cttttctgagcacagAGGGTA
Exon19( 149bp)CTGGAGgtgagggaatctggg.intron(  958bp).tctttctccccacagTGTAAT
Exon20(  86bp)GAACAGgtaacaaagggctct.intron( 2867bp).tttggctgttcccagGCCCGT
Exon21(  70bp)GCCAAGttaaagaaagtacag.intron(  801bp).cattttgttctccagTTAAAG
Exon22( 119bp)CGAAAGgtgagcaaggctgca.intron( 1434bp).ctccggtacccccagGATCTT
Exon23(1585bp)
```

FIG.4D

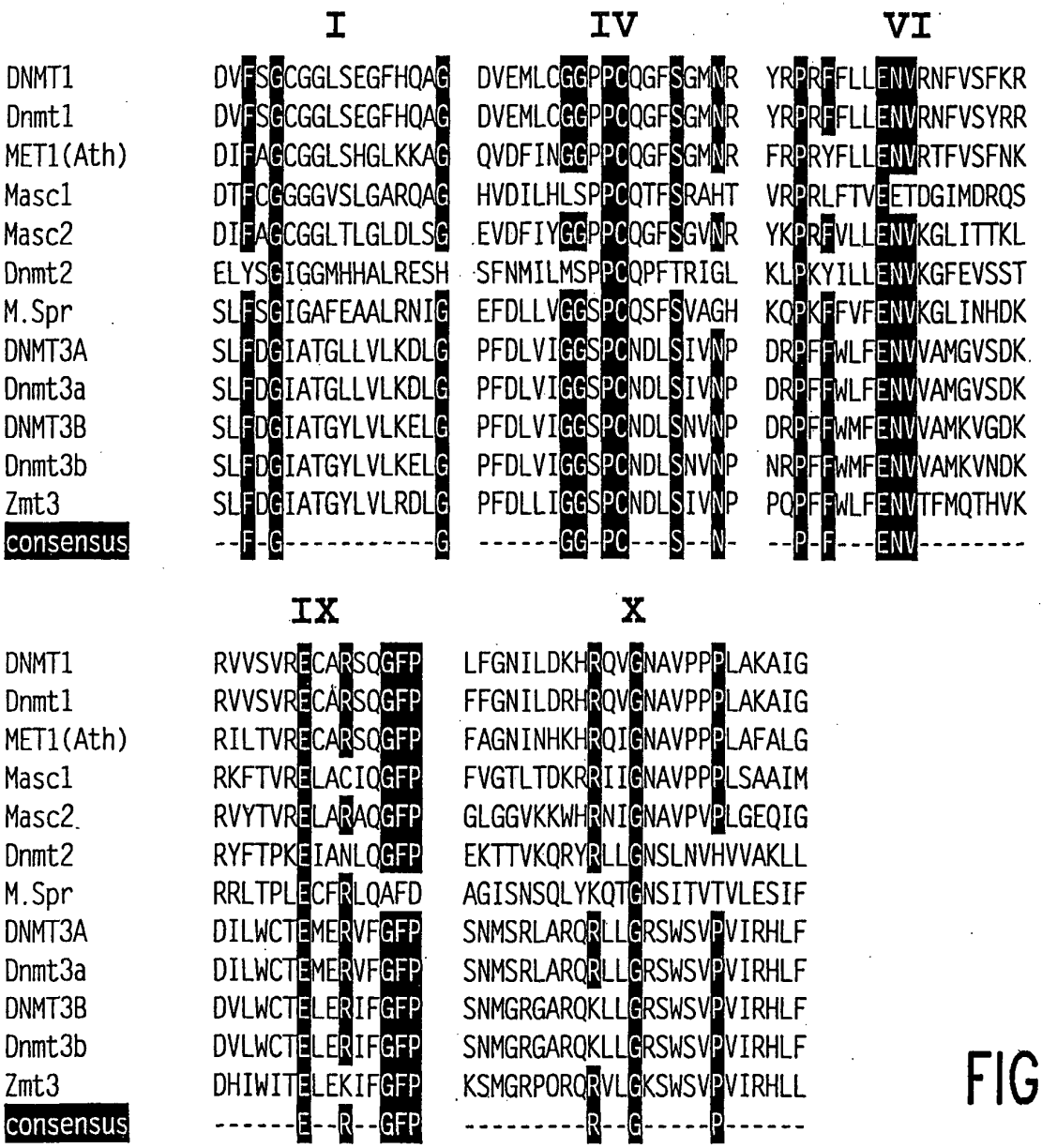

I
IV
VI
DNMT1 DVFSGCGGLSEGFHQAG DVEMLCGGPPCQGFSGMNR YRPRFFLLENVRNFVSFKR
Dnmt1 DVFSGCGGLSEGFHQAG DVEMLCGGPPCQGFSGMNR YRPRFFLLENVRNFVSYRR
MET1(Ath) DIFAGCGGLSHGLKKAG QVDFINGGPPCQGFSGMNR FRPRYFLLENVRTFVSFNK
Masc1 DTFCGGGGVSLGARQAG HVDILHLSPPCQTFSRAHT VRPRLFTVEETDGIMDRQS
Masc2 DIFAGCGGLTLGLDLSG EVDFIYGGPPCQGFSGVNR YKPRFVLLENVKGLITTKL
Dnmt2 ELYSGIGGMHHALRESH SFNMILMSPPCQPFTRIGL KLPKYILLENVKGFEVSST
M.Spr SLFSGIGAFEAALRNIG EFDLLVGGSPCQSFSVAGH KQPKFFVFENVKGLINHDK
DNMT3A SLFDGIATGLLVLKDLG PFDLVIGGSPCNDLSIVNP DRPFFWLFENVVAMGVSDK
Dnmt3a SLFDGIATGLLVLKDLG PFDLVIGGSPCNDLSIVNP DRPFFWLFENVVAMGVSDK
DNMT3B SLFDGIATGYLVLKELG PFDLVIGGSPCNDLSNVNP DRPFFWMFENVVAMKVGDK
Dnmt3b SLFDGIATGYLVLKELG PFDLVIGGSPCNDLSNVNP NRPFFWMFENVVAMKVNDK
Zmt3 SLFDGIATGYLVLRDLG PFDLLIGGSPCNDLSIVNP PQPFFWLFENVTFMQTHVK
consensus --F-G----------G ------GG-PC---S--N- --P-F---ENV--------
IX
X
DNMT1 RVVSVRECARSQGFP LFGNILDKHRQVGNAVPPPLAKAIG
Dnmt1 RVVSVRECARSQGFP FFGNILDRHRQVGNAVPPPLAKAIG
MET1(Ath) RILTVRECARSQGFP FAGNINHKHRQIGNAVPPPLAFALG
Masc1 RKFTVRELACIQGFP FVGTLTDKRRIIGNAVPPPLSAAIM
Masc2 RVYTVRELARAQGFP GLGGVKKWHRNIGNAVPVPLGEQIG
Dnmt2 RYFTPKEIANLQGFP EKTTVKQRYRLLGNSLNVHVVAKLL
M.Spr RRLTPLECFRLQAFD AGISNSQLYKQTGNSITVTVLESIF
DNMT3A DILWCTEMERVFGFP SNMSRLARQRLLGRSWSVPVIRHLF
Dnmt3a DILWCTEMERVFGFP SNMSRLARQRLLGRSWSVPVIRHLF
DNMT3B DVLWCTELERIFGFP SNMGRGARQKLLGRSWSVPVIRHLF
Dnmt3b DVLWCTELERIFGFP SNMGRGARQKLLGRSWSVPVIRHLF
Zmt3 DHIWITELEKIFGFP KSMGRPQRQRVLGKSWSVPVIRHLL
consensus ------E--R--GFP ---------R--G-----P------

FIG. 5A

```
DNMT3A          EDICISCG......SLNVTLEHPLFVGGMCQNCKNCFLECAYQYDDDGYQSYCT
Dnmt3a          EDICISCG......SLNVTLEHPFFIGGMCQNCKNCFLECAYQYDDDGYQSYCT
DNMT3B          EDGCLSCG......RKNPVSFHPLFEGGLCQTCRDRFLELFYMYDDDGYQSYCT
Dnmt3b          EDRCLSCG......KKNPVSFHPLFEGGLCQSCRDRFLELFYMYDEDGYQSYCT
Zmt3            EDFCLSCG......SMSVDIIHPLFEGKLCTNCKFNFTETLYRYDEDGYQSYCT
ATRX_Human      IVSCTACGQQVNHFQKDSIYRHPSLQVLICKNCFKYYMSDDISRDSDGMDEQCR
ATRX_Mouse      IVSCTACGQQVNHFQKDSIYRHPSLQVLICKNCFKYYMSDDISRDSDGMDEQCR
Consensus          C  C                       C  C                  C

DNMT3A          ICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKE.DPWNCYMCGHKGT
Dnmt3a          ICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIKE.DPWNCYMCGHKGT
DNMT3B          VCCEGRELLLCSNTSCCRCFCVECLEVLVGTGTAAEAKLQ.EPWSCYMCLPQRC
Dnmt3b          VCCEGRELLLCSNTSCCRCFCVECLEVLVGAGTAEDAKLQ.EPWSCYMCLPQRC
Zmt3            VCCSGMEVILCAHDSCCRSFCVDCLDILVCQGTFDRLKNV.DPWTCYLCAPETS
ATRX_Human      WCAEGGNLICC..DFCHNAFCKKCILRNLGRKELSTIMDENNQWYCYICHPEPL
ATRX_Mouse      WCAEGGNLICC..DFCHNAFCKKCILRNLGRKELSTIMDENNQWYCYICHPEPL
Consensus        C        C    C    C  C                       C  C
```

FIG.5B

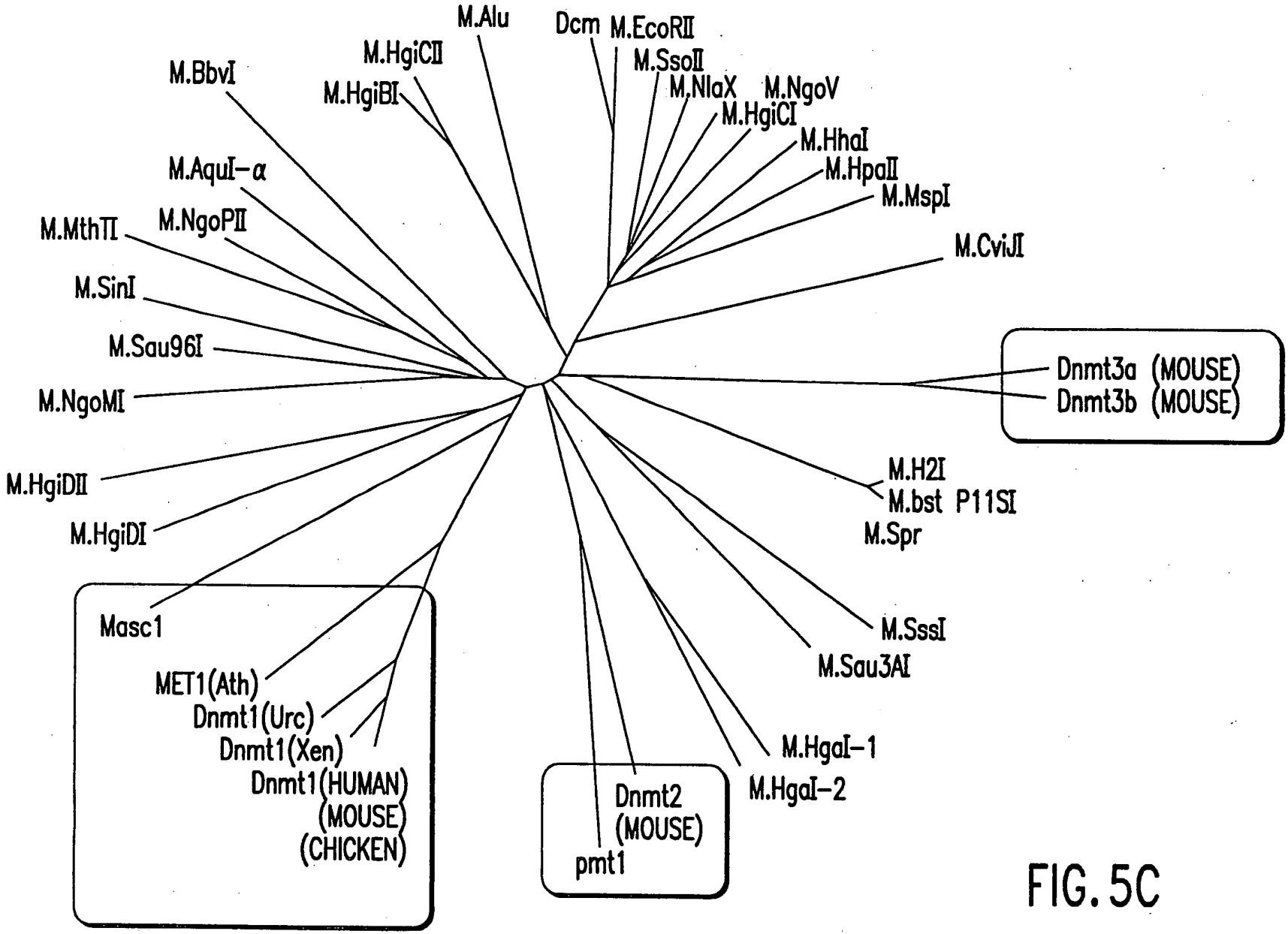

M.Alu
Dcm
M.EcoRII
M.HgiCII
M.SsoII
M.BbvI
M.HgiBI
M.NlaX
M.NgoV
M.HgiCI
M.HhaI
M.AquI-α
M.HpaII
M.MspI
M.MthTI
M.NgoPII
M.CviJI
M.SinI
M.Sau96I
Dnmt3a (MOUSE)
Dnmt3b (MOUSE)
M.NgoMI
M.H2I
M.bst P11SI
M.HgiDII
M.Spr
M.HgiDI
M.SssI
Masc1
M.Sau3AI
MET1(Ath)
Dnmt1(Urc)
M.HgaI-1
Dnmt1(Xen)
Dnmt1(HUMAN)
(MOUSE)
(CHICKEN)
Dnmt2
(MOUSE)
M.HgaI-2
pmt1

FIG. 5C

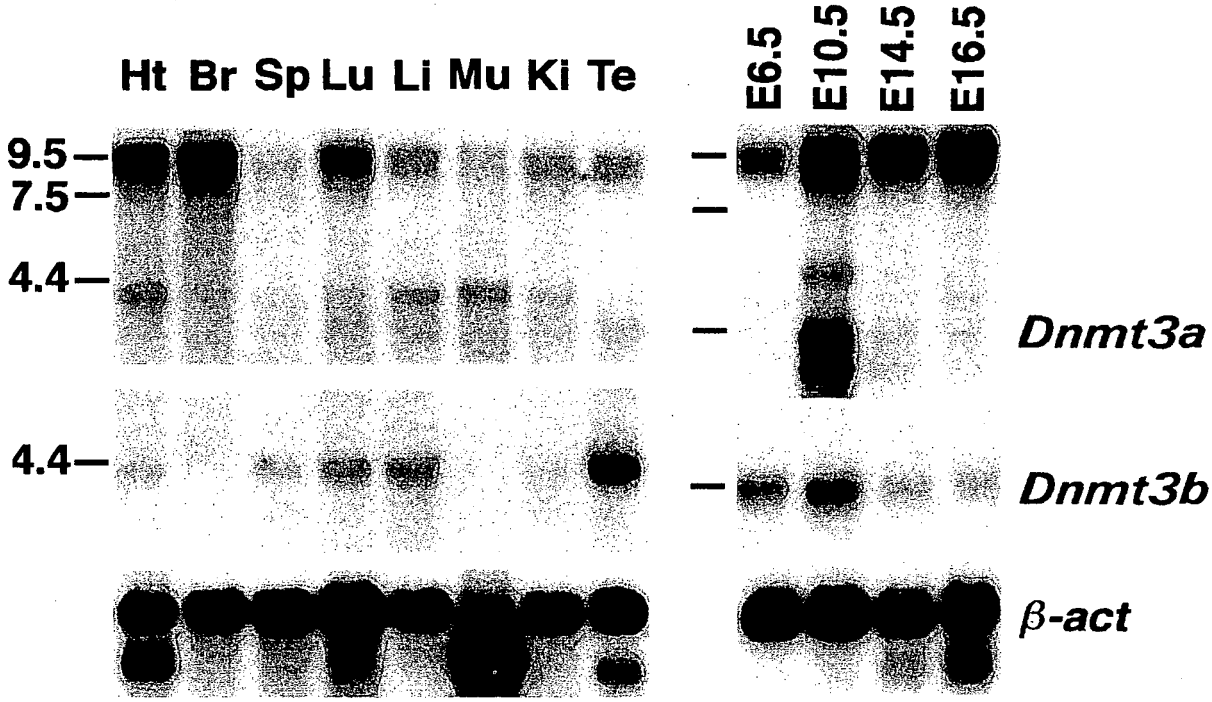


FIG.6A

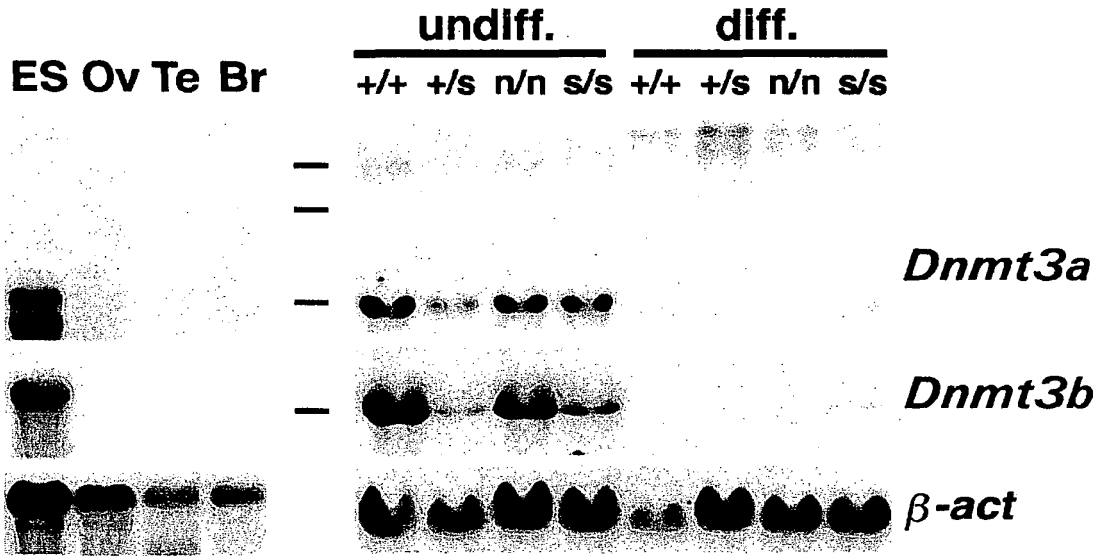


FIG.6B

FIG.6C

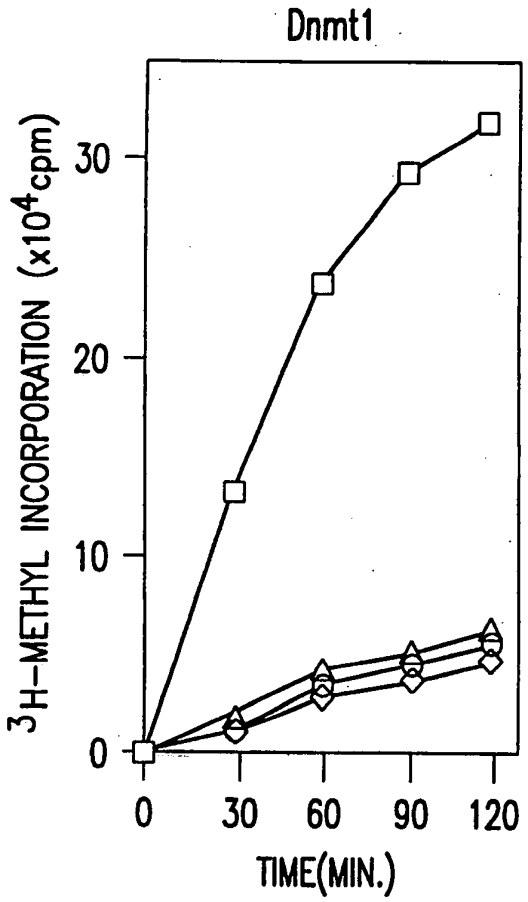


FIG. 7A

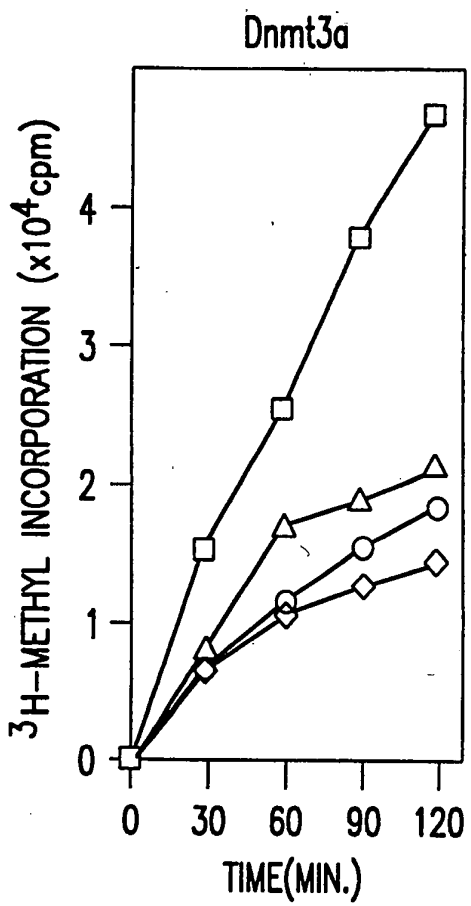


FIG. 7B

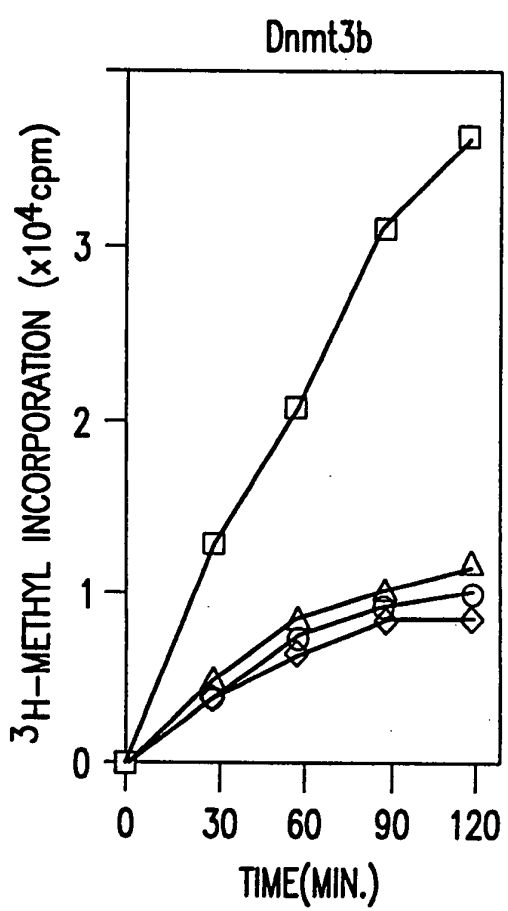


FIG. 7C

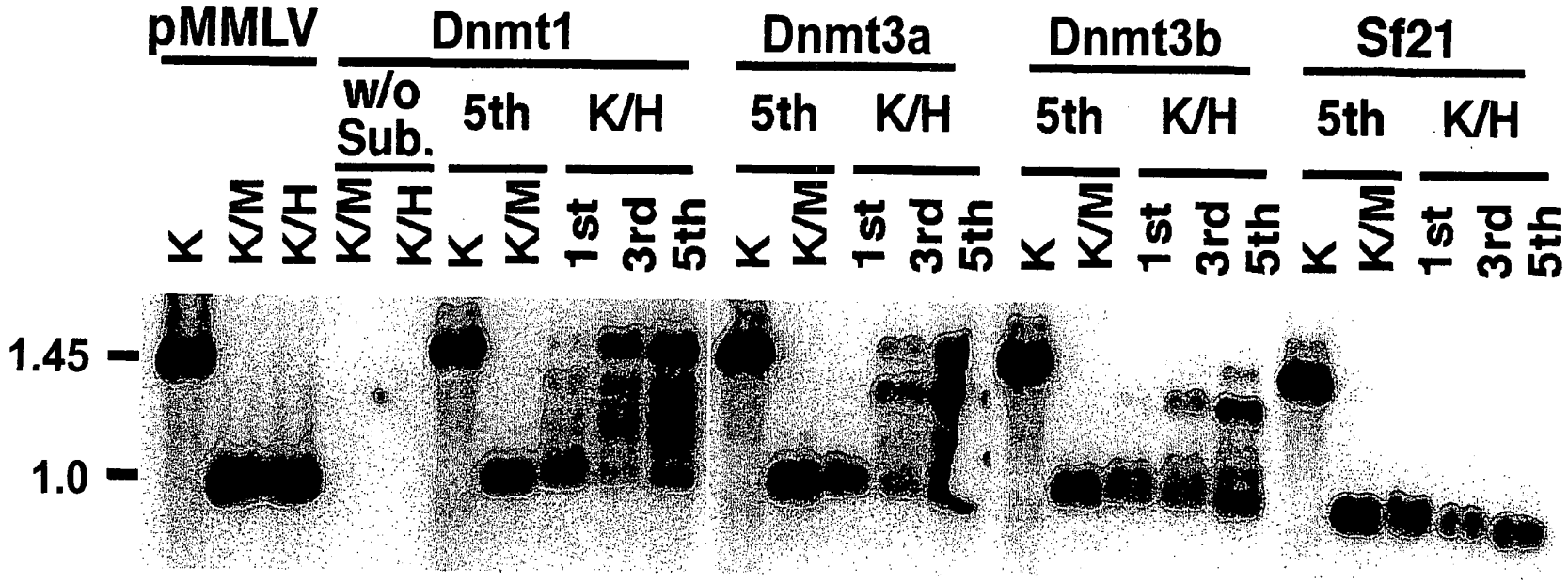

pMMLV
Dnmt1
w/o Sub.
5th
K/H
Dnmt3a
5th
K/H
Dnmt3b
5th
K/H
Sf21
5th
K/H
K
K/M
K/H
K/M
K/H
K
K/M
1st
3rd
5th
K
K/M
1st
3rd
5th
K
K/M
1st
3rd
5th
K
K/M
1st
3rd
5th
1.45
1.0

FIG. 7D

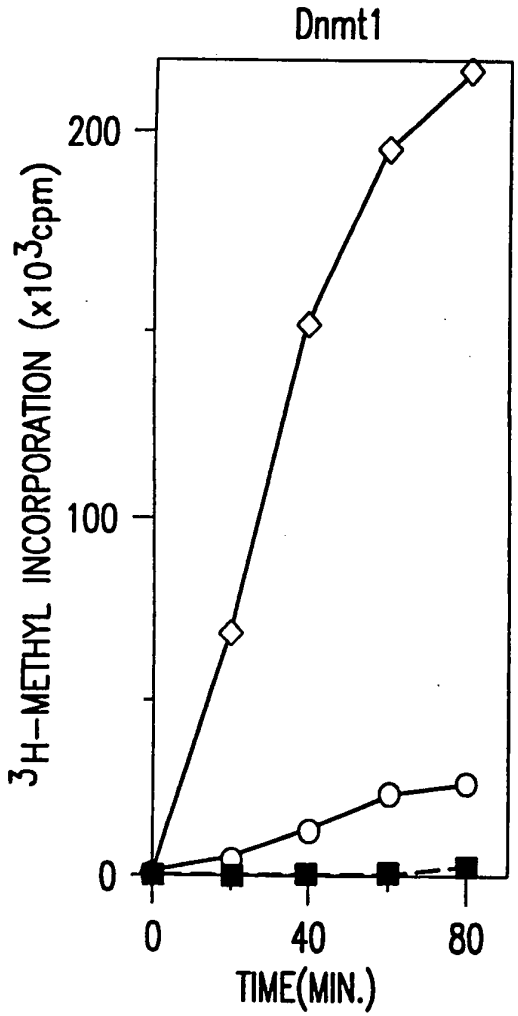


FIG. 8A

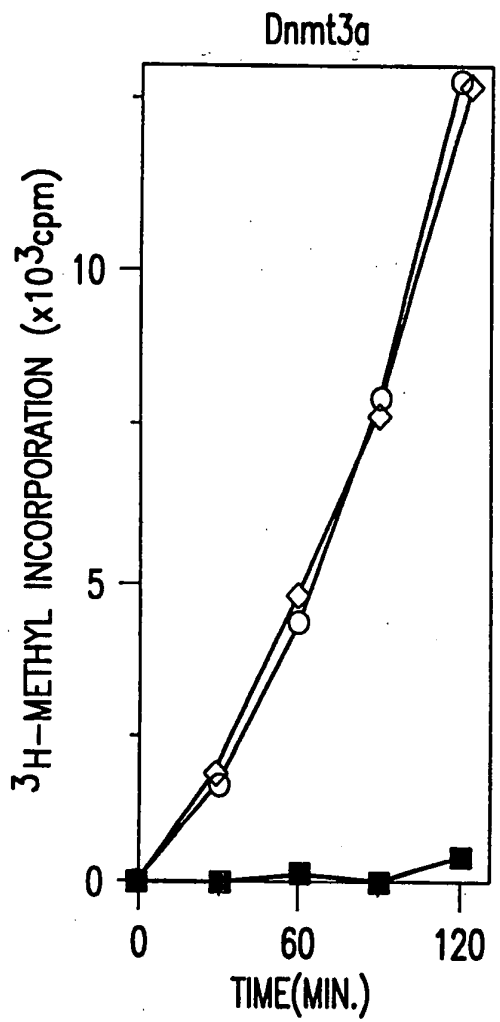


FIG. 8B

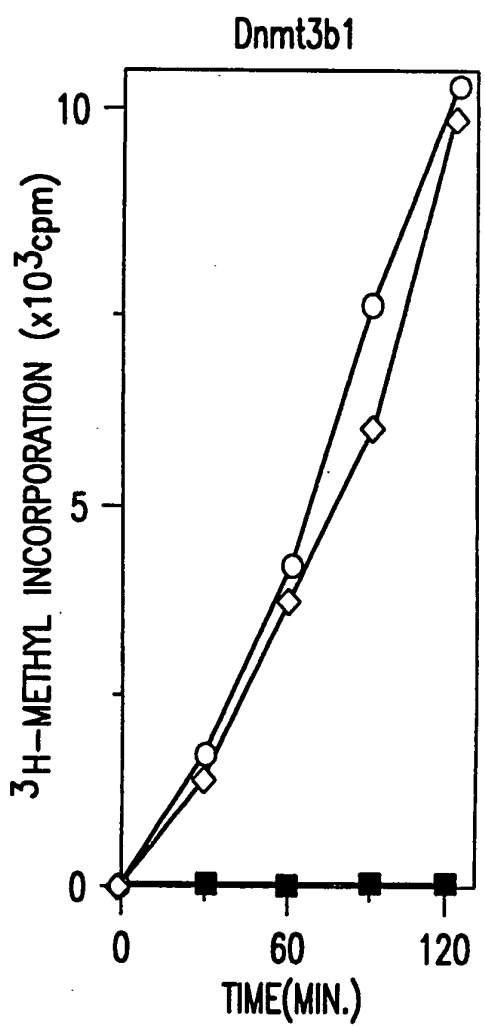


FIG. 8C

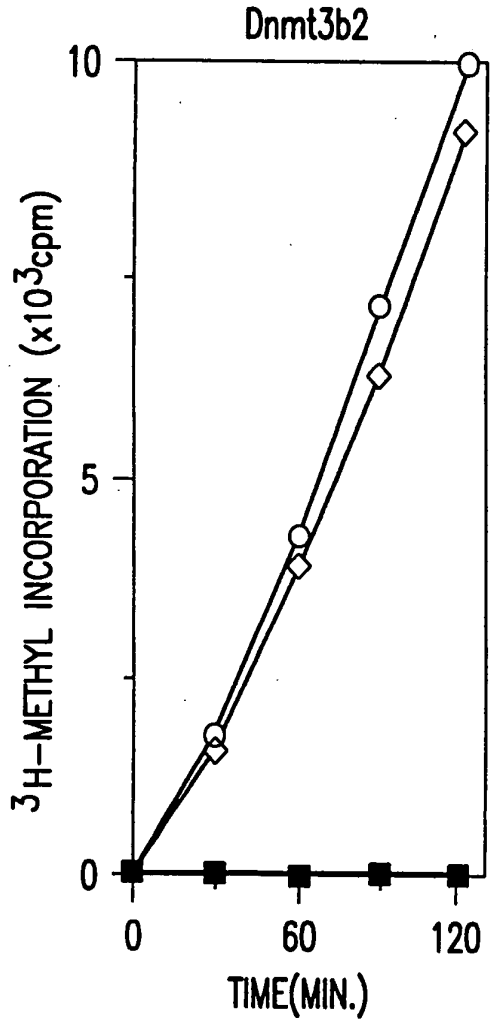

Dnmt3b2
10
5
0
0
60
120
TIME(MIN.)
^{3}H-METHYL INCORPORATION (x10^{3}cpm)

FIG. 8D

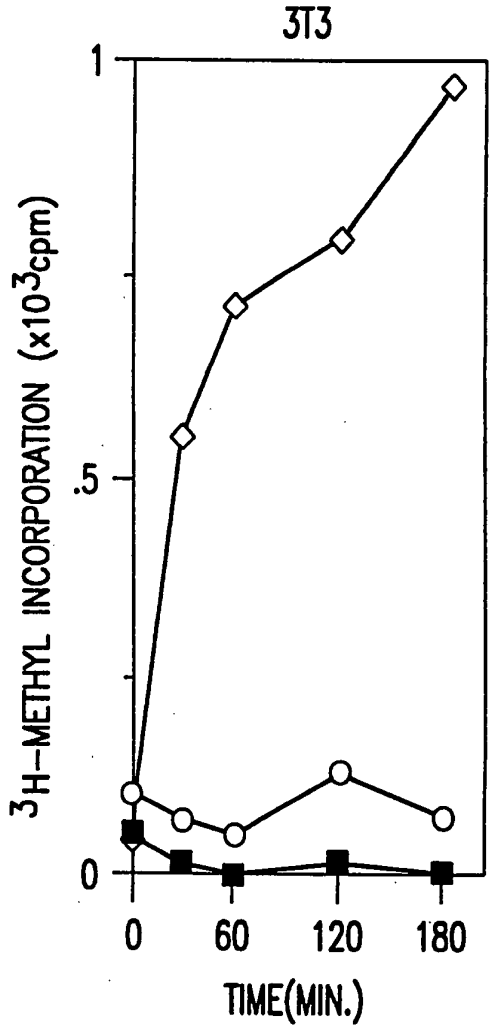
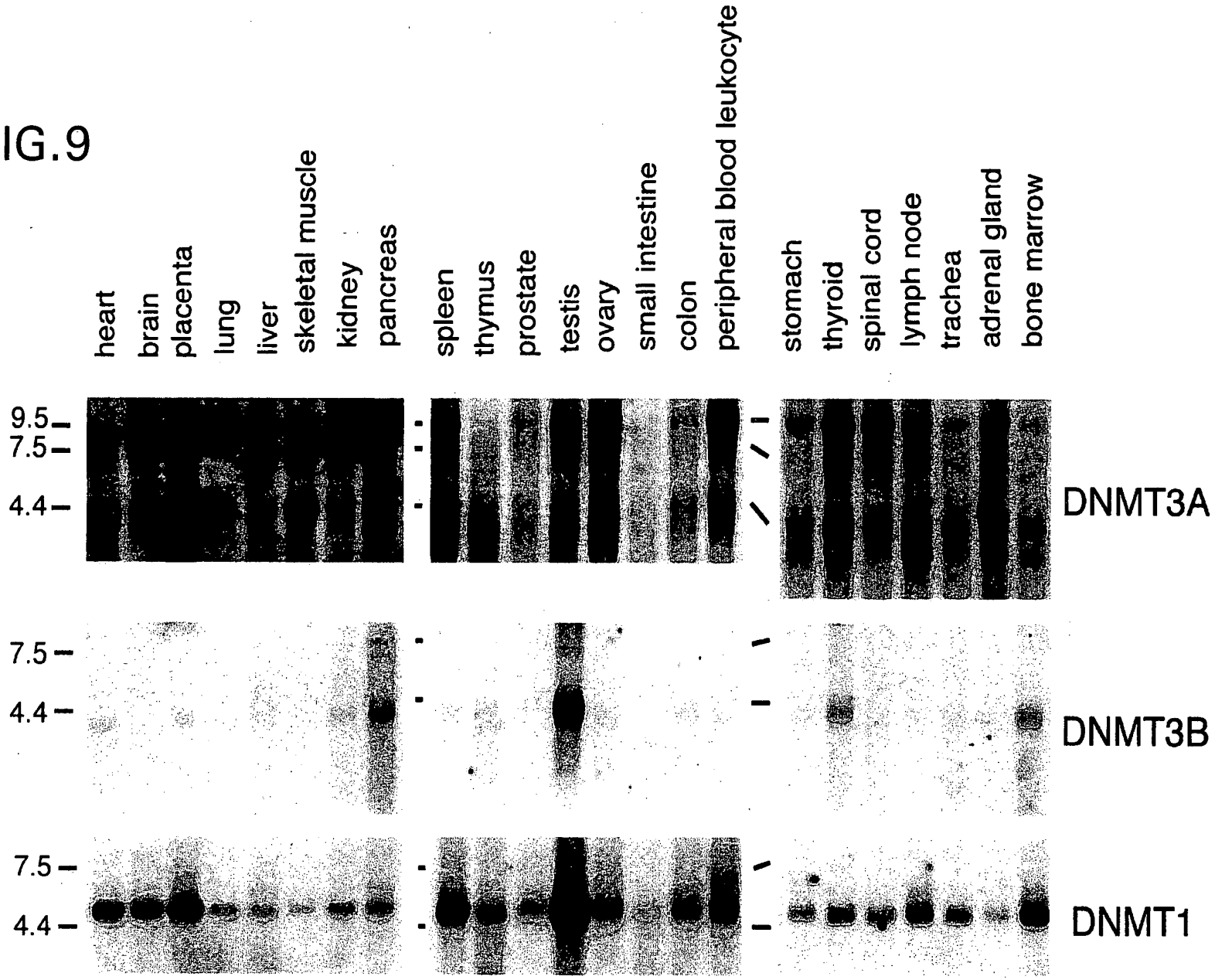

3T3
1
.5
0
0
60
120
180
TIME(MIN.)
^{3}H-METHYL INCORPORATION (x10^{3}cpm)

FIG. 8E

FIG.9


heart
brain
placenta
lung
liver
skeletal muscle
kidney
pancreas
spleen
thymus
prostate
testis
ovary
small intestine
colon
peripheral blood leukocyte
stomach
thyroid
spinal cord
lymph node
trachea
adrenal gland
bone marrow
9.5
7.5
4.4
DNMT3A
7.5
4.4
DNMT3B
7.5
4.4
DNMT1

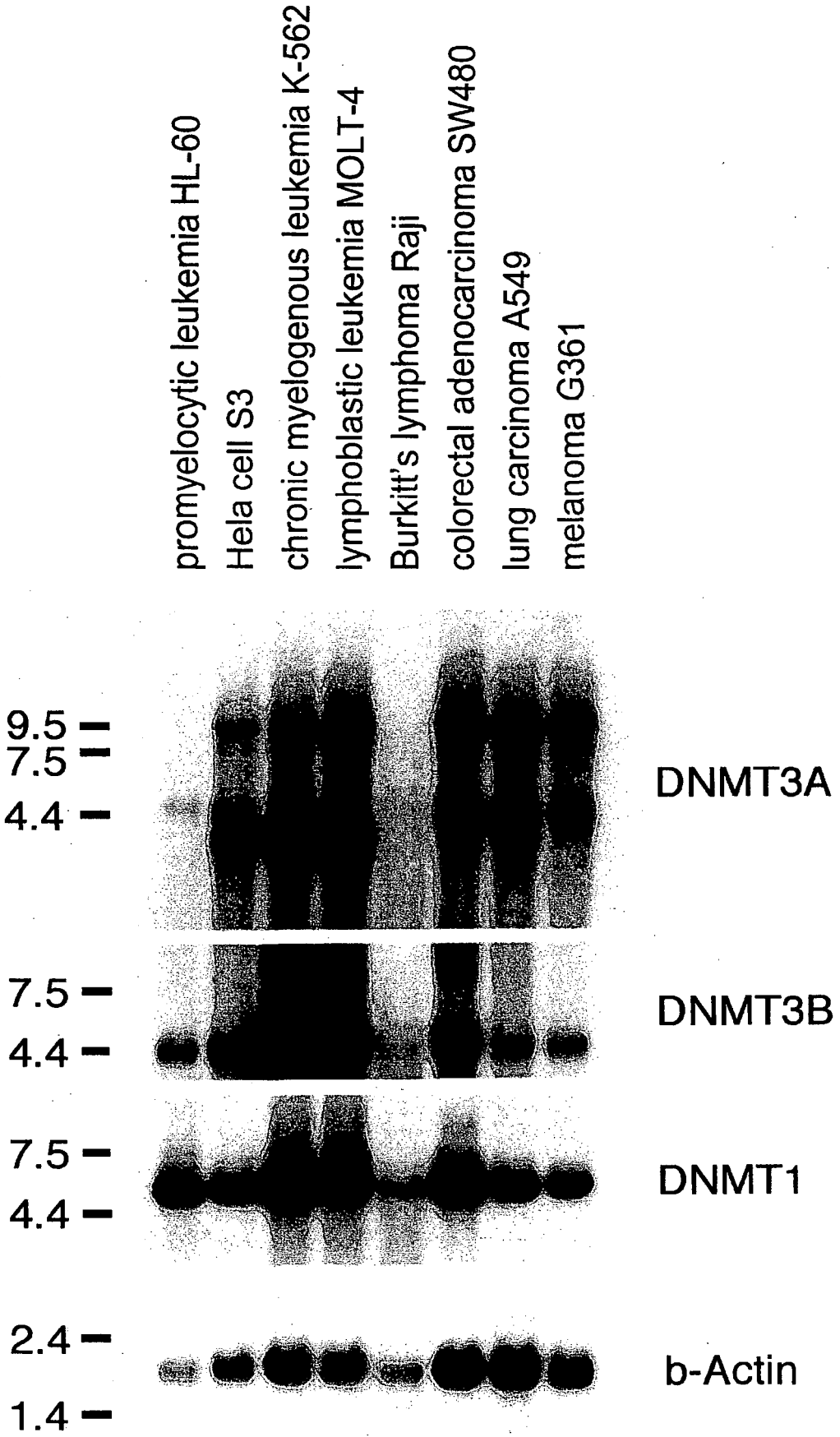

promyelocytic leukemia HL-60
Hela cell S3
chronic myelogenous leukemia K-562
lymphoblastic leukemia MOLT-4
Burkitt's lymphoma Raji
colorectal adenocarcinoma SW480
lung carcinoma A549
melanoma G361
9.5
7.5
4.4
DNMT3A
7.5
4.4
DNMT3B
7.5
4.4
DNMT1
2.4
1.4
b-Actin

FIG. 10

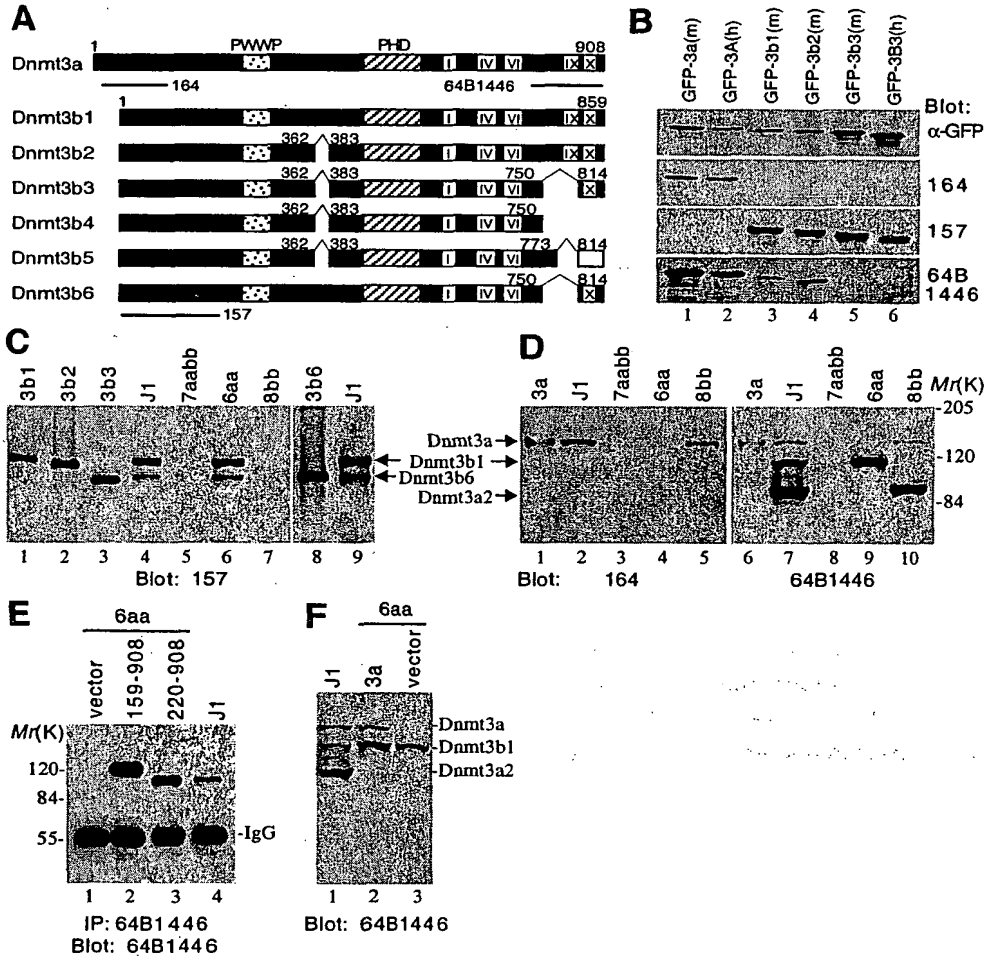

A
Dnmt3a
Dnmt3b1
Dnmt3b2
Dnmt3b3
Dnmt3b4
Dnmt3b5
Dnmt3b6
PWWP
PHD
164
64B1446
157
B
GFP-3a(m)
GFP-3A(h)
GFP-3b1(m)
GFP-3b2(m)
GFP-3b3(m)
GFP-3B3(h)
Blot: α-GFP
164
157
64B 1446
C
Dnmt3a
Dnmt3b1
Dnmt3b6
Dnmt3a2
Blot: 157
D
Mr(K)
205
120
84
Blot: 164
64B1446
E
6aa
vector
159-908
220-908
J1
Mr(K)
120
84
55
IgG
IP: 64B1446
Blot: 64B1446
F
6aa
J1
3a
vector
Dnmt3a
Dnmt3b1
Dnmt3a2
Blot: 64B1446

FIG. 11

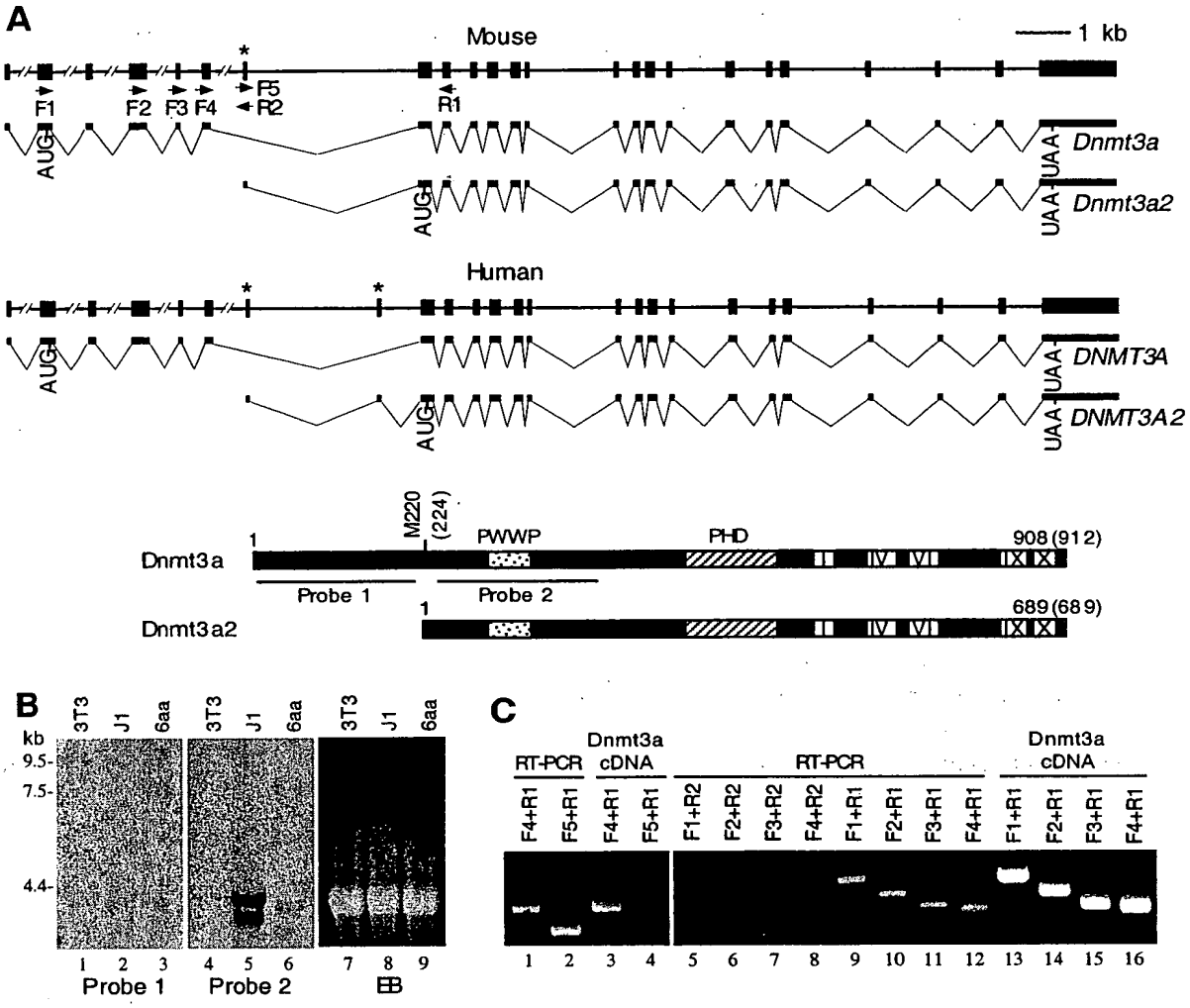

A
Mouse
1 kb
F1
F2
F3
F4
F5
R2
R1
AUG
UAA
Dnmt3a
Dnmt3a2
Human
DNMT3A
DNMT3A2
M220
(224)
PWWP
PHD
908 (912)
Dnmt3a
Probe 1
Probe 2
689 (689)
Dnmt3a2
B
kb
9.5-
7.5-
4.4-
3T3
J1
6aa
1 2 3 4 5 6 7 8 9
Probe 1
Probe 2
EB
C
RT-PCR
Dnmt3a cDNA
F4+R1
F5+R1
F1+R2
F2+R2
F3+R2
F4+R2
F1+R1
F2+R1
F3+R1
F4+R1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16

FIG. 12

Mouse Dnmt3a2 cDNA sequence:

```
   1 ccgcccccaaccccaacgccccctgcccctcccccagacgggcagctatttacagagc
  60 ttcgggccggggctcacacctgagctgtactgcagaggggctgcacctggccttatggg
 119 ctgagaagaaagccaaggtaattgcagtaatgaatgctgtggaagagaaccaggcctct
 178 ggagagtctcagaaggtggaggaggccagccctcctgctgtgcagcagcccacggaccc
 237 tgcttctccgactgtggccaccacccctgagccagtaggaggggatgctggggacaaga
 296 atgctaccaaagcagccgacgatgagcctgagtatgaggatggccggggctttggcatt
 355 ggagagctggtgtgggggaaacttcggggcttctcctggtggccaggccgaattgtgtc
 414 ttggtggatgacaggccggagccgagcagctgaaggcactcgctgggtcatgtggttcg
 473 gagatggcaagttctcagtggtgtgtgtggagaagctcatgccgctgagctccttctgc
 532 agtgcattccaccaggccacctacaacaagcagcccatgtaccgcaaagccatctacga
 591 agtcctccaggtggccagcagccgtgccgggaagctgtttccagcttgccatgacagtg
 650 atgaaagtgacagtggcaaggctgtggaagtgcagaacaagcagatgattgaatgggcc
 709 ctcggtggcttccagccctcgggtcctaagggcctggagccaccagaagaagagaagaa
 768 tccttacaaggaagtttacaccgacatgtgggtggagcctgaagcagctgcttacgccc
 827 cacccccaccagccaagaaacccagaaagagcacaacagagaaacctaaggtcaaggag
 886 atcattgatgagcgcacaagggagcggctggtgtatgaggtgcgccagaagtgcagaaa
 945 catcgaggacatttgtatctcatgtgggagcctcaatgtcaccctggagcacccactct
1004 tcattggaggcatgtgccagaactgtaagaactgcttcttggagtgtgcttaccagtat
1063 gacgacgatgggtaccagtcctattgcaccatctgctgtgggggcgtgaagtgctcat
1122 gtgtgggaacaacaactgctgcaggtgcttttgtgtcgagtgtgtggatctcttggtgg
1181 ggccaggagctgctcaggcagccattaaggaagacccctggaactgctacatgtgcggg
1240 cataagggcacctatgggctgctgcgaagacgggaagactggccttctcgactccagat
1299 gttctttgccaataaccatgaccaggaatttgacccccaaaggtttacccacctgtgc
1358 cagctgagaagaggaagcccatccgcgtgctgtctctctttgatgggattgctacaggg
1417 ctcctggtgctgaaggacctgggcatccaagtggaccgctacattgcctccgaggtgtg
1476 tgaggactccatcacggtgggcatggtgcggcaccagggaaagatcatgtacgtcgggg
1535 acgtccgcagcgtcacacagaagcatatccaggagtggggcccattcgacctggtgatt
1594 ggaggcagtccctgcaatgacctctccattgtcaaccctgcccgcaagggactttatga
1653 gggtactggccgcctcttctttgagttctaccgcctcctgcatgatgcgcggcccaagg
1712 agggagatgatcgccccttcttctggctctttgagaatgtggtggccatgggcgttagt
1771 gacaagagggacatctcgcgatttcttgagtctaaccccgtgatgattgacgccaaaga
1830 agtgtctgctgcacacagggcccgttacttctggggtaaccttcctggcatgaacaggc
1889 ctttggcatccactgtgaatgataagctggagctgcaagagtgtctggagcacggcaga
1948 atagccaagttcagcaaagtgaggaccattaccaccaggtcaaactctataaagcaggg
2007 caaagaccagcatttccccgtcttcatgaacgagaaggaggacatcctgtggtgcactg
2066 aaatggaaagggtgtttggcttccccgtccactacacagacgtctccaacatgagccgc
2125 ttggcgaggcagagactgctgggccgatcgtggagcgtgccggtcatccgccacctctt
2184 cgctccgctgaaggaatattttgcttgtgtgtaagggacatgggggcaaactgaagtag
2243 tgatgataaaaaagttaaacaaacaaacaaacaaaaaacaaaacaaaacaataaaacac
2302 caagaacgagaaaaaaa
```

FIG. 13A

Mouse Dnmt3a2 amino acid sequence:

```
   1 MNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKNATKAADDEP
  60 EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCV
 119 EKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPACHDSDESDSGKAVE
 178 VQNKQMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRK
 237 STTEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGGMCQNCK
 296 NCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIK
 355 EDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRV
 414 LSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHI
 473 QEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL
 532 FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKL
 591 ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPV
 650 HYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV
```

FIG. 13B

Human DNMT3A2 cDNA sequence:

```
   1 ccgcccccagccccatcgccccttcccctcccccaagacgggcagctacttccagagc
  60 ttcagggccgcggctcacacctgagcgcgactgcagaggggctgcacctggccttatgg
 119 ggatcctggagcgggttgtgagaaggaatgggcgcgtggatcgtagcctgaaagacgag
 178 tgtgatacggctgagaagaaagccaaggtcattgcaggaatgaatgctgtggaagaaaa
 237 ccaggggcccggggagtctcagaaggtggaggaggccagccctcctgctgtgcagcagc
 296 ccactgaccccgcatcccccactgtggctaccacgcctgagcccgtggggtccgatgct
 355 ggggacaagaatgccaccaaagcaggcgatgacgagccagagtacgaggacggccgggg
 414 ctttggcattggggagctggtgtgggggaaactgcggggcttctcctggtggccaggcc
 473 gcattgtgtcttggtggatgacgggccggagccgagcagctgaaggcacccgctgggtc
 532 atgtggttcggagacggcaaattctcagtggtgtgttgagaagctgatgccgctgag
 591 ctcgttttgcagtgcgttccaccaggccacgtacaacaagcagcccatgtaccgcaaag
 650 ccatctacgaggtcctgcaggtggccagcagccgcgcggggaagctgttcccggtgtgc
 709 cacgacagcgatgagagtgacactgccaaggccgtggaggtgcagaacaagcccatgat
 768 tgaatgggccctgggggcttccagccttctggccctaagggcctggagccaccagaag
 827 aagagaagaatccctacaaagaagtgtacacggacatgtgggtggaacctgaggcagct
 886 gcctacgcaccacctccaccagccaaaaagccccggaagagcacagcggagaagcccaa
 945 ggtcaaggagattattgatgagcgcacaagagagcggctggtgtacgaggtgcggcaga
1004 agtgccggaacattgaggacatctgcatctcctgtgggagcctcaatgttaccctggaa
1063 caccccctcttcgttggaggaatgtgccaaaactgcaagaactgctttctggagtgtgc
1122 gtaccagtacgacgacgacggctaccagtcctactgcaccatctgctgtgggggccgtg
1181 aggtgctcatgtgcggaaacaacaactgctgcaggtgcttttgcgtggagtgtgtggac
1240 ctcttggtggggccgggggctgcccaggcagccattaaggaagacccctggaactgcta
1299 catgtgcgggcacaagggtacctacgggctgctgcggcggcgagaggactggccctccc
1358 ggctccagatgttcttcgctaataaccacgaccaggaatttgaccctccaaaggtttac
1417 ccacctgtcccagctgagaagaggaagcccatccgggtgctgtctctctttgatggaat
1476 cgctacagggctcctggtgctgaaggacttgggcattcaggtggaccgctacattgcct
1535 cggaggtgtgtgaggactccatcacggtgggcatggtgcggcaccaggggaagatcatg
1594 tacgtcggggacgtccgcagcgtcacacagaagcatatccaggagtggggcccattcga
1653 tctggtgattgggggcagtccctgcaatgacctctccatcgtcaaccctgctcgcaagg
1712 gcctctacgagggcactggccggctcttctttgagttctaccgcctcctgcatgatgcg
1771 cggcccaaggagggagatgatcgccccttcttctggctctttgagaatgtggtggccat
1830 gggcgttagtgacaagagggacatctcgcgatttctcgagtccaaccctgtgatgattg
1889 atgccaaagaagtgtcagctgcacacagggcccgctacttctggggtaaccttcccggt
1948 atgaacaggccgttggcatccactgtgaatgataagctggagctgcaggagtgtctgga
2007 gcatggcaggatagccaagttcagcaaagtgaggaccattactacgaggtcaaactcca
2066 taaagcagggcaaagaccagcattttcctgtcttcatgaatgagaaagaggacatctta
2125 tggtgcactgaaatggaaagggtatttggtttcccagtccactatactgacgtctccaa
2184 catgagccgcttggcgaggcagagactgctgggccggtcatggagcgtgccagtcatcc
2243 gccacctcttcgctccgctgaaggagtattttgcgtgtgtgtaagggacatgggggcaa
2302 actgaggtagcgacacaaagttaaacaaacaaacaaaaaacacaaaacataataaaaca
2361 ccaagaacatg
```

FIG. 13C

Human DNMT3A2 amino acid sequence:

```
1   MNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKNATKAGDDEP
60  EYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAEGTRWVMWFGDGKFSVVCV
119 EKLMPLSSFCSAFHQATYNKQPMYRKAIYEVLQVASSRAGKLFPVCHDSDESDTAKAVE
178 VQNKPMIEWALGGFQPSGPKGLEPPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRK
237 STAEKPKVKEIIDERTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFVGGMCQNCK
296 NCFLECAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQAAIK
355 EDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVYPPVPAEKRKPIRV
414 LSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVGMVRHQGKIMYVGDVRSVTQKHI
473 QEWGPFDLVIGGSPCNDLSIVNPARKGLYEGTGRLFFEFYRLLHDARPKEGDDRPFFWL
532 FENVVAMGVSDKRDISRFLESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKL
591 ELQECLEHGRIAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPV
650 HYTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV
```

FIG. 13D

```
                        10         20         30         40         50
Dnmt3a2               1 ccgcccccaa ccccaacgcc ccctgcccct ccccccagac gggcagctat 50
DNMT3A2               1 ccgcccccag ccccatcgcc cccttccccc cccccaagac gggcagctac 50

                       60         70         80         90        100
Dnmt3a2              51 ttacagagct tc-gggccgg ggctcacacc tgagctgtac tgcagagggg 100
DNMT3A2              51 ttccagagct tcagggccgc ggctcacacc tgagcgcgac tgcagagggg 100

                      110        120        130        140        150
Dnmt3a2             101 ctgcacctgg ccttatgg-- ---------- ---------- ---------- 150
DNMT3A2             101 ctgcacctgg ccttatgggg atcctggagc gggttgtgag aaggaatggg 150

                      160        170        180        190        200
Dnmt3a2             151 ---------- ---------- ---------- ------gctg agaagaaagc 200
DNMT3A2             151 cgcgtggatc gtagcctgaa agacgagtgt gatacggctg agaagaaagc 200

                      210        220        230        240        250
Dnmt3a2             201 caaggtaatt gcagtaatga atgctgtgga agagaaccag gcctctggag 250
DNMT3A2             201 caaggtcatt gcaggaatga atgctgtgga agaaaaccag gggcccgggg 250

                      260        270        280        290        300
Dnmt3a2             251 agtctcagaa ggtggaggag gccagccctc ctgctgtgca gcagcccacg 300
DNMT3A2             251 agtctcagaa ggtggaggag gccagccctc ctgctgtgca gcagcccact 300

                      310        320        330        340        350
Dnmt3a2             301 gaccctgctt ctccgactgt ggccaccacc cctgagccag taggagggga 350
DNMT3A2             301 gaccccgcat cccccactgt ggctaccacg cctgagcccg tggggtccga 350

                      360        370        380        390        400
Dnmt3a2             351 tgctggggac aagaatgcta ccaaagcagc cgacgatgag cctgagtatg 400
DNMT3A2             351 tgctggggac aagaatgcca ccaaagcagg cgatgacgag ccagagtacg 400

                      410        420        430        440        450
Dnmt3a2             401 aggatggccg gggctttggc attggagagc tggtgtgggg gaaacttcgg 450
DNMT3A2             401 aggacggccg gggctttggc attggggagc tggtgtgggg gaaactgcgg 450

                      460        470        480        490        500
Dnmt3a2             451 ggcttctcct ggtggccagg ccgaattgtg tcttggtgga tgacaggccg 500
DNMT3A2             451 ggcttctcct ggtggccagg ccgcattgtg tcttggtgga tgacgggccg 500

                      510        520        530        540        550
Dnmt3a2             501 gagccgagca gctgaaggca ctcgctgggt catgtggttc ggagatggca 550
DNMT3A2             501 gagccgagca gctgaaggca cccgctgggt catgtggttc ggagacggca 550

                      560        570        580        590        600
Dnmt3a2             551 agttctcagt ggtgtgtgtg gagaagctca tgccgctgag ctccttctgc 600
DNMT3A2             551 aattctcagt ggtgtgtgtt gagaagctga tgccgctgag ctcgttttgc 600

                      610        620        630        640        650
Dnmt3a2             601 agtgcattcc accaggccac ctacaacaag cagcccatgt accgcaaagc 650
DNMT3A2             601 agtgcgttcc accaggccac gtacaacaag cagcccatgt accgcaaagc 650

                      660        670        680        690        700
Dnmt3a2             651 catctacgaa gtcctccagg tggccagcag ccgtgccggg aagctgtttc 700
DNMT3A2             651 catctacgag gtcctgcagg tggccagcag ccgcgcgggg aagctgttcc 700
```

FIG. 13E-1

```
                    710         720         730         740         750
Dnmt3a2      701 cagcttgcca tgacagtgat gaaagtgaca gtggcaaggc tgtggaagtg 750
DNMT3A2      701 cggtgtgcca cgacagcgat gagagtgaca ctgccaaggc cgtggaggtg 750

                    760         770         780         790         800
Dnmt3a2      751 cagaacaagc agatgattga atgggccctc ggtggcttcc agccctcggg 800
DNMT3A2      751 cagaacaagc ccatgattga atgggccctg gggggcttcc agccttctgg 800

                    810         820         830         840         850
Dnmt3a2      801 tcctaagggc ctggagccac cagaagaaga gaagaatcct tacaaggaag 850
DNMT3A2      801 ccctaagggc ctggagccac cagaagaaga gaagaatccc tacaaagaag 850

                    860         870         880         890         900
Dnmt3a2      851 tttacaccga catgtgggtg gagcctgaag cagctgctta cgccccaccc 900
DNMT3A2      851 tgtacacgga catgtgggtg gaacctgagg cagctgccta cgcaccacct 900

                    910         920         930         940         950
Dnmt3a2      901 ccaccagcca agaaacccag aaagagcaca acagagaaac ctaaggtcaa 950
DNMT3A2      901 ccaccagcca aaaagccccg gaagagcaca gcggagaagc ccaaggtcaa 950

                    960         970         980         990        1000
Dnmt3a2      951 ggagatcatt gatgagcgca caagggagcg gctggtgtat gaggtgcgcc 1000
DNMT3A2      951 ggagattatt gatgagcgca caagagagcg gctggtgtac gaggtgcggc 1000

                   1010        1020        1030        1040        1050
Dnmt3a2     1001 agaagtgcag aaacatcgag gacatttgta tctcatgtgg gagcctcaat 1050
DNMT3A2     1001 agaagtgccg gaacattgag gacatctgca tctcctgtgg gagcctcaat 1050

                   1060        1070        1080        1090        1100
Dnmt3a2     1051 gtcaccctgg agcacccact cttcattgga ggcatgtgcc agaactgtaa 1100
DNMT3A2     1051 gttaccctgg aacacccccc cttcgttgga ggaatgtgcc aaaactgcaa 1100

                   1110        1120        1130        1140        1150
Dnmt3a2     1101 gaactgcttc ttggagtgtg cttaccagta tgacgacgat gggtaccagt 1150
DNMT3A2     1101 gaactgcttt ctggagtgtg cgtaccagta cgacgacgac ggctaccagt 1150

                   1160        1170        1180        1190        1200
Dnmt3a2     1151 cctattgcac catctgctgt gggggcgtg aagtgctcat gtgtgggaac 1200
DNMT3A2     1151 cctactgcac catctgctgt gggggccgtg aggtgctcat gtgcggaaac 1200

                   1210        1220        1230        1240        1250
Dnmt3a2     1201 aacaactgct gcaggtgctt ttgtgtcgag tgtgtggatc tcttggtggg 1250
DNMT3A2     1201 aacaactgct gcaggtgctt ttgcgtggag tgtgtggacc tcttggtggg 1250

                   1260        1270        1280        1290        1300
Dnmt3a2     1251 gccaggagct gctcaggcag ccattaagga agacccctgg aactgctaca 1300
DNMT3A2     1251 gccgggggct gcccaggcag ccattaagga agacccctgg aactgctaca 1300

                   1310        1320        1330        1340        1350
Dnmt3a2     1301 tgtgcgggca taagggcacc tatgggctgc tgcgaagacg ggaagactgg 1350
DNMT3A2     1301 tgtgcgggca caagggtacc tacgggctgc tgcggcggcg agaggactgg 1350

                   1360        1370        1380        1390        1400
Dnmt3a2     1351 ccttctcgac tccagatgtt ctttgccaat aaccatgacc aggaatttga 1400
DNMT3A2     1351 ccctcccggc tccagatgtt cttcgctaat aaccacgacc aggaatttga 1400
```

FIG. 13E-2

```
                    1410        1420        1430        1440        1450
Dnmt3a2       1401 ccccccaaag gtttacccac ctgtgccagc tgagaagagg aagcccatcc 1450
DNMT3A2       1401 ccctccaaag gtttacccac ctgtcccagc tgagaagagg aagcccatcc 1450

                    1460        1470        1480        1490        1500
Dnmt3a2       1451 gcgtgctgtc tctctttgat gggattgcta cagggctcct ggtgctgaag 1500
DNMT3A2       1451 gggtgctgtc tctctttgat ggaatcgcta cagggctcct ggtgctgaag 1500

                    1510        1520        1530        1540        1550
Dnmt3a2       1501 gacctgggca tccaagtgga ccgctacatt gcctccgagg tgtgtgagga 1550
DNMT3A2       1501 gacttgggca ttcaggtgga ccgctacatt gcctcggagg tgtgtgagga 1550

                    1560        1570        1580        1590        1600
Dnmt3a2       1551 ctccatcacg gtgggcatgg tgcggcacca gggaaagatc atgtacgtcg 1600
DNMT3A2       1551 ctccatcacg gtgggcatgg tgcggcacca ggggaagatc atgtacgtcg 1600

                    1610        1620        1630        1640        1650
Dnmt3a2       1601 gggacgtccg cagcgtcaca cagaagcata tccaggagtg gggcccattc 1650
DNMT3A2       1601 gggacgtccg cagcgtcaca cagaagcata tccaggagtg gggcccattc 1650

                    1660        1670        1680        1690        1700
Dnmt3a2       1651 gacctggtga ttggaggcag tccctgcaat gacctctcca ttgtcaaccc 1700
DNMT3A2       1651 gatctggtga ttgggggcag tccctgcaat gacctctcca tcgtcaaccc 1700

                    1710        1720        1730        1740        1750
Dnmt3a2       1701 tgcccgcaag ggactttatg agggtactgg ccgcctcttc tttgagttct 1750
DNMT3A2       1701 tgctcgcaag ggcctctacg agggcactgg ccggctcttc tttgagttct 1750

                    1760        1770        1780        1790        1800
Dnmt3a2       1751 accgcctcct gcatgatgcg cggcccaagg agggagatga tcgccccttc 1800
DNMT3A2       1751 accgcctcct gcatgatgcg cggcccaagg agggagatga tcgccccttc 1800

                    1810        1820        1830        1840        1850
Dnmt3a2       1801 ttctggctct ttgagaatgt ggtggccatg ggcgttagtg acaagaggga 1850
DNMT3A2       1801 ttctggctct ttgagaatgt ggtggccatg ggcgttagtg acaagaggga 1850

                    1860        1870        1880        1890        1900
Dnmt3a2       1851 catctcgcga tttcttgagt ctaaccccgt gatgattgac gccaaagaag 1900
DNMT3A2       1851 catctcgcga tttctcgagt ccaaccctgt gatgattgat gccaaagaag 1900

                    1910        1920        1930        1940        1950
Dnmt3a2       1901 tgtctgctgc acacagggcc cgttacttct ggggtaacct tcctggcatg 1950
DNMT3A2       1901 tgtcagctgc acacagggcc cgctacttct ggggtaacct tcccggtatg 1950

                    1960        1970        1980        1990        2000
Dnmt3a2       1951 aacaggcctt tggcatccac tgtgaatgat aagctggagc tgcaagagtg 2000
DNMT3A2       1951 aacaggccgt tggcatccac tgtgaatgat aagctggagc tgcaggagtg 2000

                    2010        2020        2030        2040        2050
Dnmt3a2       2001 tctggagcac ggcagaatag ccaagttcag caaagtgagg accattacca 2050
DNMT3A2       2001 tctggagcat ggcaggatag ccaagttcag caaagtgagg accattacta 2050

                    2060        2070        2080        2090        2100
Dnmt3a2       2051 ccaggtcaaa ctctataaag cagggcaaag accagcattt ccccgtcttc 2100
DNMT3A2       2051 cgaggtcaaa ctccataaag cagggcaaag accagcattt tcctgtcttc 2100
```

FIG. 13E-3

```
                    2110        2120        2130        2140        2150
Dnmt3a2     2101 atgaacgaga aggaggacat cctgtggtgc actgaaatgg aaagggtgtt 2150
DNMT3A2     2101 atgaatgaga aagaggacat cttatggtgc actgaaatgg aaagggtatt 2150

                2160        2170        2180        2190        2200
Dnmt3a2     2151 tggcttcccc gtccactaca cagacgtctc caacatgagc cgcttggcga 2200
DNMT3A2     2151 tggtttccca gtccactata ctgacgtctc caacatgagc cgcttggcga 2200

                2210        2220        2230        2240        2250
Dnmt3a2     2201 ggcagagact gctgggccga tcgtggagcg tgccggtcat ccgccacctc 2250
DNMT3A2     2201 ggcagagact gctgggccgg tcatggagcg tgccagtcat ccgccacctc 2250

                2260        2270        2280        2290        2300
Dnmt3a2     2251 ttcgctccgc tgaaggaata ttttgcttgt gtgtaaggga catgggggca 2300
DNMT3A2     2251 ttcgctccgc tgaaggagta ttttgcgtgt gtgtaaggga catgggggca 2300

                2310        2320        2330        2340        2350
Dnmt3a2     2301 aactgaagta gtgatgataa aaaagttaaa caaacaaaca aacaaaaaac 2350
DNMT3A2     2301 aactgaggta gcgac----a caaagttaaa caaacaaac- ---aaaaaac 2350

                2360        2370        2380
Dnmt3a2     2351 aaaacaaaac aataaaacac caagaacgag
DNMT3A2     2351 acaaaacat- aataaaacac caagaacatg
```

FIG. 13E-4

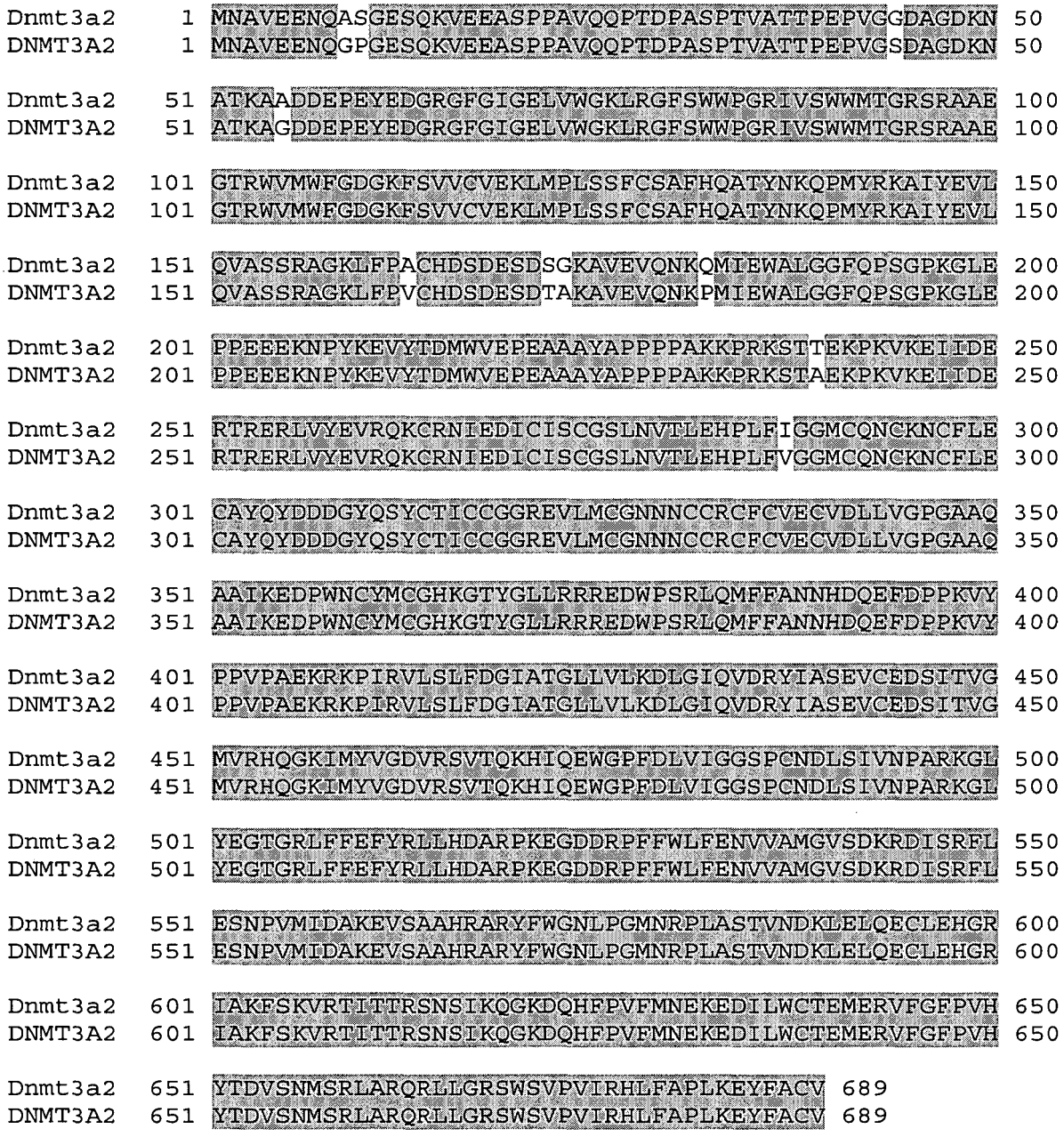

Dnmt3a2 1 MNAVEENQASGESQKVEEASPPAVQQPTDPASPTVATTPEPVGGDAGDKN 50
DNMT3A2 1 MNAVEENQGPGESQKVEEASPPAVQQPTDPASPTVATTPEPVGSDAGDKN 50
Dnmt3a2 51 ATKAADDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAE 100
DNMT3A2 51 ATKAGDDEPEYEDGRGFGIGELVWGKLRGFSWWPGRIVSWWMTGRSRAAE 100
Dnmt3a2 101 GTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVL 150
DNMT3A2 101 GTRWVMWFGDGKFSVVCVEKLMPLSSFCSAFHQATYNKQPMYRKAIYEVL 150
Dnmt3a2 151 QVASSRAGKLFPACHDSDESDSGKAVEVQNKQMIEWALGGFQPSGPKGLE 200
DNMT3A2 151 QVASSRAGKLFPVCHDSDESDTAKAVEVQNKPMIEWALGGFQPSGPKGLE 200
Dnmt3a2 201 PPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRKSTTEKPKVKEIIDE 250
DNMT3A2 201 PPEEEKNPYKEVYTDMWVEPEAAAYAPPPPAKKPRKSTAEKPKVKEIIDE 250
Dnmt3a2 251 RTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFIGGMCQNCKNCFLE 300
DNMT3A2 251 RTRERLVYEVRQKCRNIEDICISCGSLNVTLEHPLFVGGMCQNCKNCFLE 300
Dnmt3a2 301 CAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQ 350
DNMT3A2 301 CAYQYDDDGYQSYCTICCGGREVLMCGNNNCCRCFCVECVDLLVGPGAAQ 350
Dnmt3a2 351 AAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVY 400
DNMT3A2 351 AAIKEDPWNCYMCGHKGTYGLLRRREDWPSRLQMFFANNHDQEFDPPKVY 400
Dnmt3a2 401 PPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG 450
DNMT3A2 401 PPVPAEKRKPIRVLSLFDGIATGLLVLKDLGIQVDRYIASEVCEDSITVG 450
Dnmt3a2 451 MVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGL 500
DNMT3A2 451 MVRHQGKIMYVGDVRSVTQKHIQEWGPFDLVIGGSPCNDLSIVNPARKGL 500
Dnmt3a2 501 YEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFL 550
DNMT3A2 501 YEGTGRLFFEFYRLLHDARPKEGDDRPFFWLFENVVAMGVSDKRDISRFL 550
Dnmt3a2 551 ESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGR 600
DNMT3A2 551 ESNPVMIDAKEVSAAHRARYFWGNLPGMNRPLASTVNDKLELQECLEHGR 600
Dnmt3a2 601 IAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVH 650
DNMT3A2 601 IAKFSKVRTITTRSNSIKQGKDQHFPVFMNEKEDILWCTEMERVFGFPVH 650
Dnmt3a2 651 YTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV 689
DNMT3A2 651 YTDVSNMSRLARQRLLGRSWSVPVIRHLFAPLKEYFACV 689

FIG. 13F

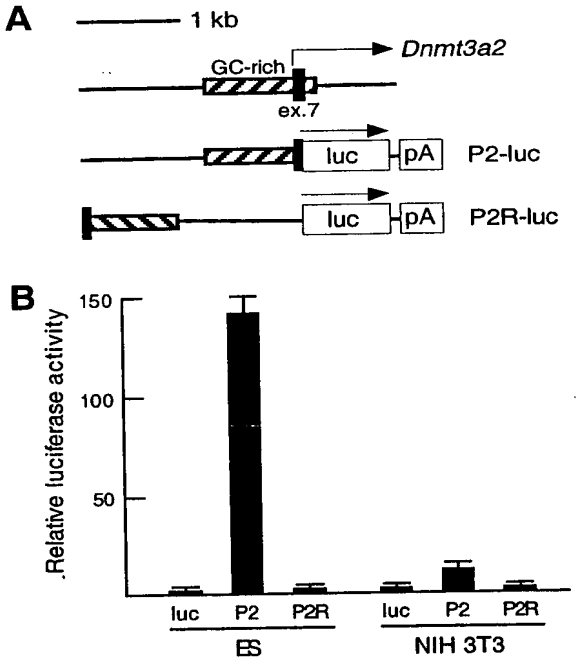

A
1 kb
Dnmt3a2
GC-rich
ex.7
luc
pA
P2-luc
luc
pA
P2R-luc
B
Relative luciferase activity
150
100
50
luc
P2
P2R
luc
P2
P2R
ES
NIH 3T3

FIG. 14

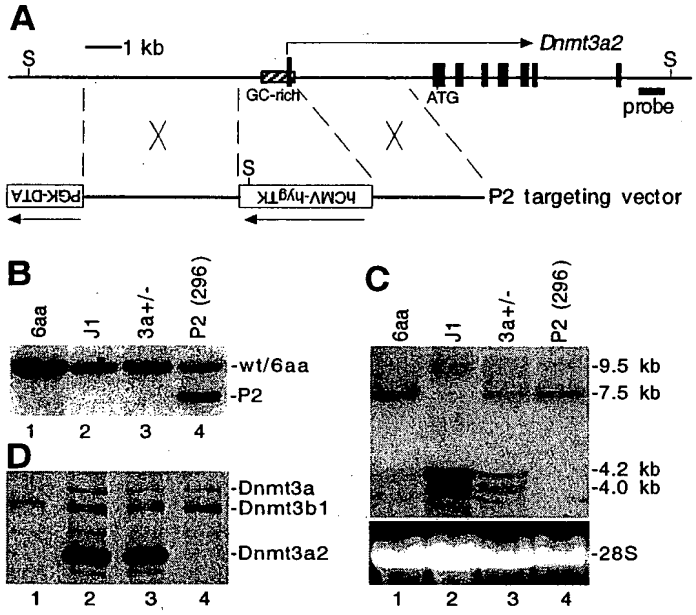

A
1 kb
S
Dnmt3a2
S
GC-rich
ATG
probe
S
PGK-DTA
hCMV-hygTK
P2 targeting vector
B
6aa
J1
3a+/-
P2 (296)
-wt/6aa
-P2
1 2 3 4
C
6aa
J1
3a+/-
P2 (296)
-9.5 kb
-7.5 kb
-4.2 kb
-4.0 kb
-28S
1 2 3 4
D
-Dnmt3a
-Dnmt3b1
-Dnmt3a2
1 2 3 4

FIG. 15

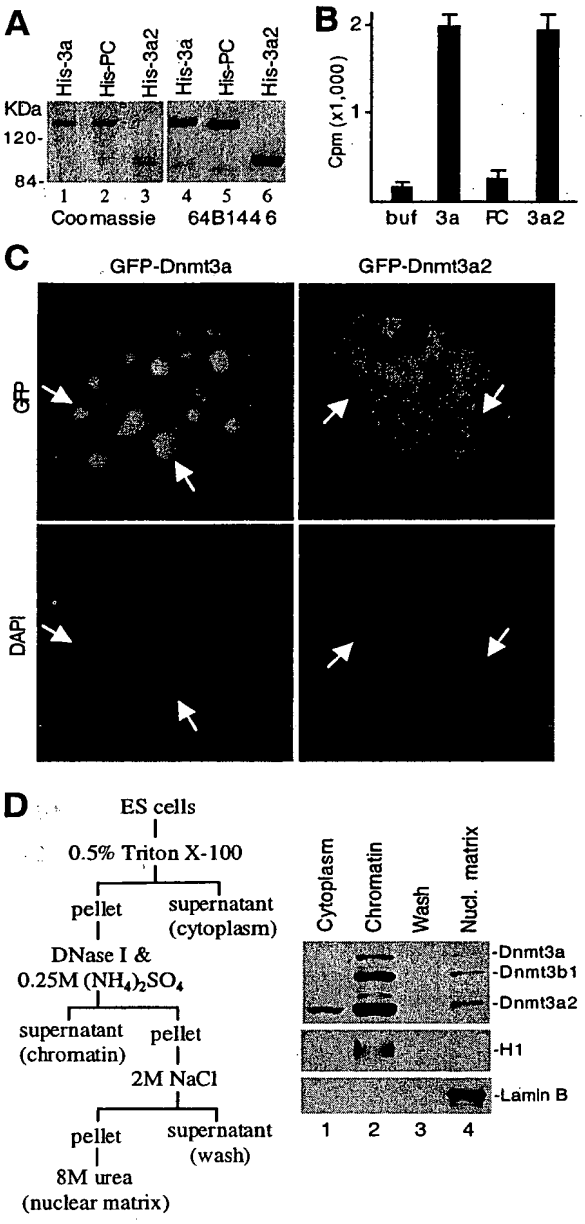

A
His-3a
His-PC
His-3a2
His-3a
His-PC
His-3a2
KDa
120-
84-
1 2 3 4 5 6
Coomassie
64B1446
B
Cpm (x1,000)
2
1
buf
3a
FC
3a2
C
GFP-Dnmt3a
GFP-Dnmt3a2
GFP
DAPI
D
ES cells
0.5% Triton X-100
pellet
supernatant
(cytoplasm)
DNase I &
0.25M (NH4)2SO4
supernatant
(chromatin)
pellet
2M NaCl
pellet
supernatant
(wash)
8M urea
(nuclear matrix)
Cytoplasm
Chromatin
Wash
Nucl. matrix
-Dnmt3a
-Dnmt3b1
-Dnmt3a2
-H1
-Lamin B
1 2 3 4

FIG. 16

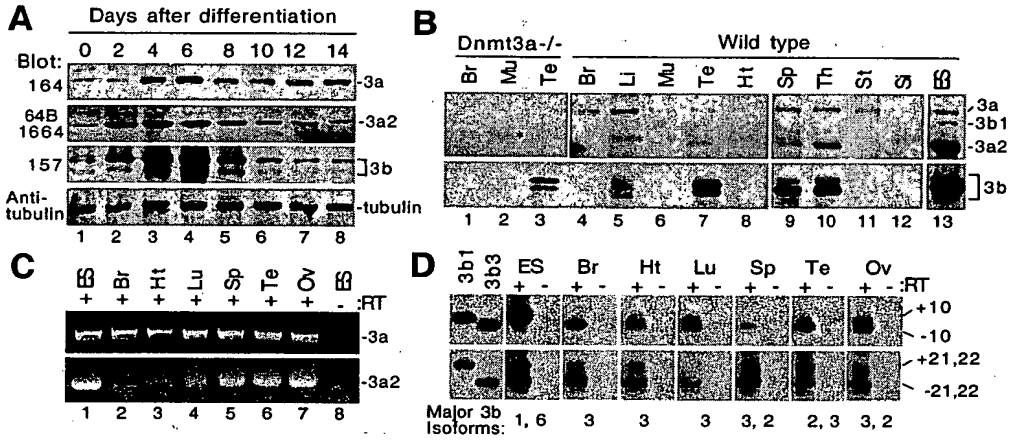

A
Days after differentiation
0 2 4 6 8 10 12 14
Blot:
164
-3a
64B
1664
-3a2
157
3b
Anti-
tubulin
-tubulin
1 2 3 4 5 6 7 8
B
Dnmt3a-/-
Wild type
Br Mu Te Br Li Mu Te Ht Sp Th St Si ES
-3a
-3b1
-3a2
3b
1 2 3 4 5 6 7 8 9 10 11 12 13
C
ES Br Ht Lu Sp Te Ov ES
+ + + + + + + - :RT
-3a
-3a2
1 2 3 4 5 6 7 8
D
3b1
3b3
ES Br Ht Lu Sp Te Ov
+ - + - + - + - + - + - + - :RT
+10
-10
+21,22
-21,22
Major 3b
Isoforms: 1, 6 3 3 3 3, 2 2, 3 3, 2

FIG. 17

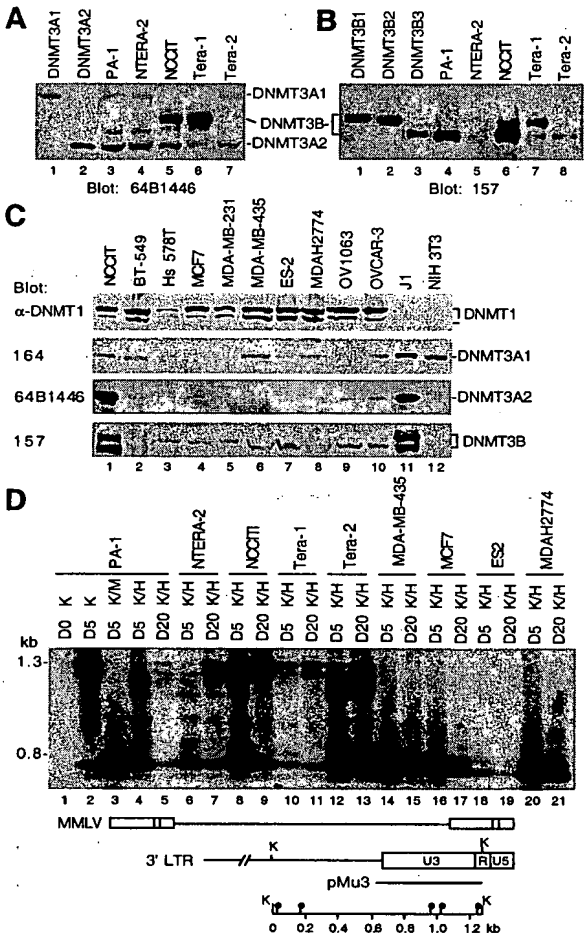

A
DNMT3A1
DNMT3A2
PA-1
NTERA-2
NCCIT
Tera-1
Tera-2
DNMT3A1
DNMT3B
DNMT3A2
1 2 3 4 5 6 7
Blot: 64B1446
B
DNMT3B1
DNMT3B2
DNMT3B3
PA-1
NTERA-2
NCCIT
Tera-1
Tera-2
1 2 3 4 5 6 7 8
Blot: 157
C
Blot:
NCCIT
BT-549
Hs 578T
MCF7
MDA-MB-231
MDA-MB-435
ES-2
MDAH2774
OV1063
OVCAR-3
J1
NIH 3T3
α-DNMT1
DNMT1
164
DNMT3A1
64B1446
DNMT3A2
157
DNMT3B
1 2 3 4 5 6 7 8 9 10 11 12
D
PA-1
NTERA-2
NCCIT
Tera-1
Tera-2
MDA-MB-435
MCF7
ES2
MDAH2774
D0 K
D5 K
D5 K/M
D5 K/H
D20 K/H
kb
1.3
0.8
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
MMLV
3' LTR
U3
R
U5
pMu3
0 0.2 0.4 0.6 0.8 1.0 1.2 kb

FIG. 18

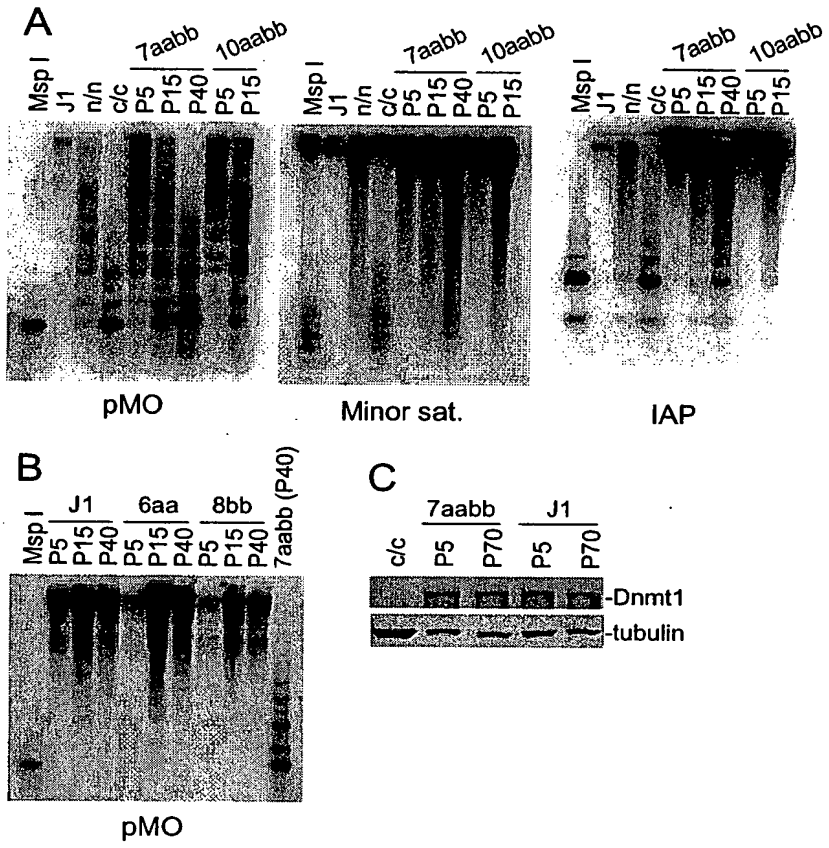


FIG. 19

0302-66.vsd/1

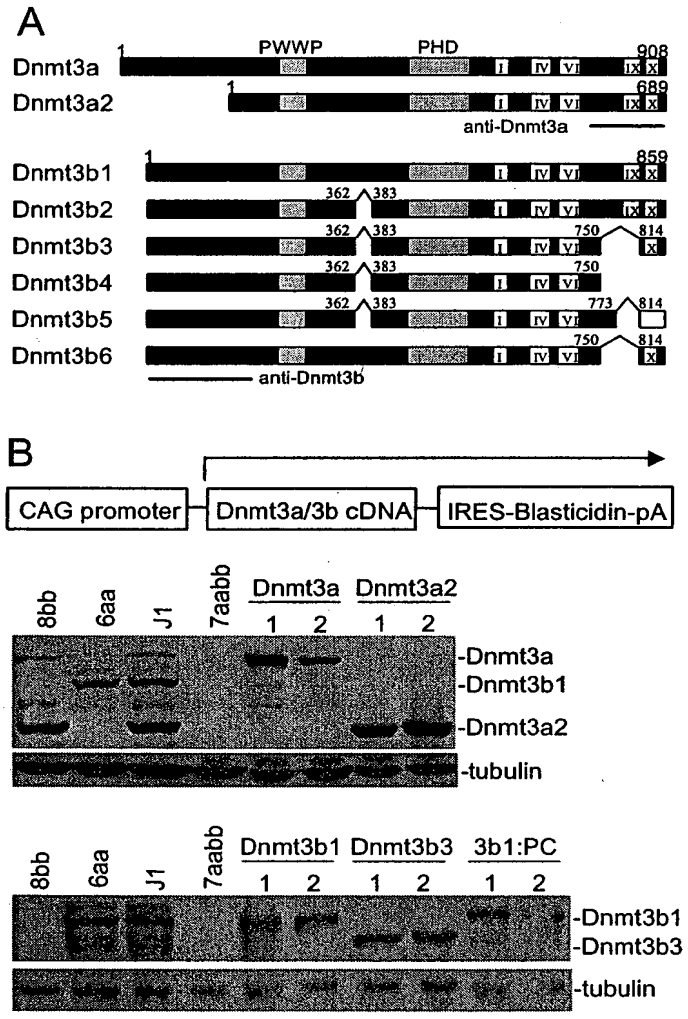

A
Dnmt3a
Dnmt3a2
PWWP
PHD
anti-Dnmt3a
Dnmt3b1
Dnmt3b2
Dnmt3b3
Dnmt3b4
Dnmt3b5
Dnmt3b6
anti-Dnmt3b
B
CAG promoter
Dnmt3a/3b cDNA
IRES-Blasticidin-pA
8bb
6aa
J1
7aabb
Dnmt3a
Dnmt3a2
-Dnmt3a
-Dnmt3b1
-Dnmt3a2
-tubulin
Dnmt3b1
Dnmt3b3
3b1:PC
-Dnmt3b1
-Dnmt3b3
-tubulin

FIG. 20

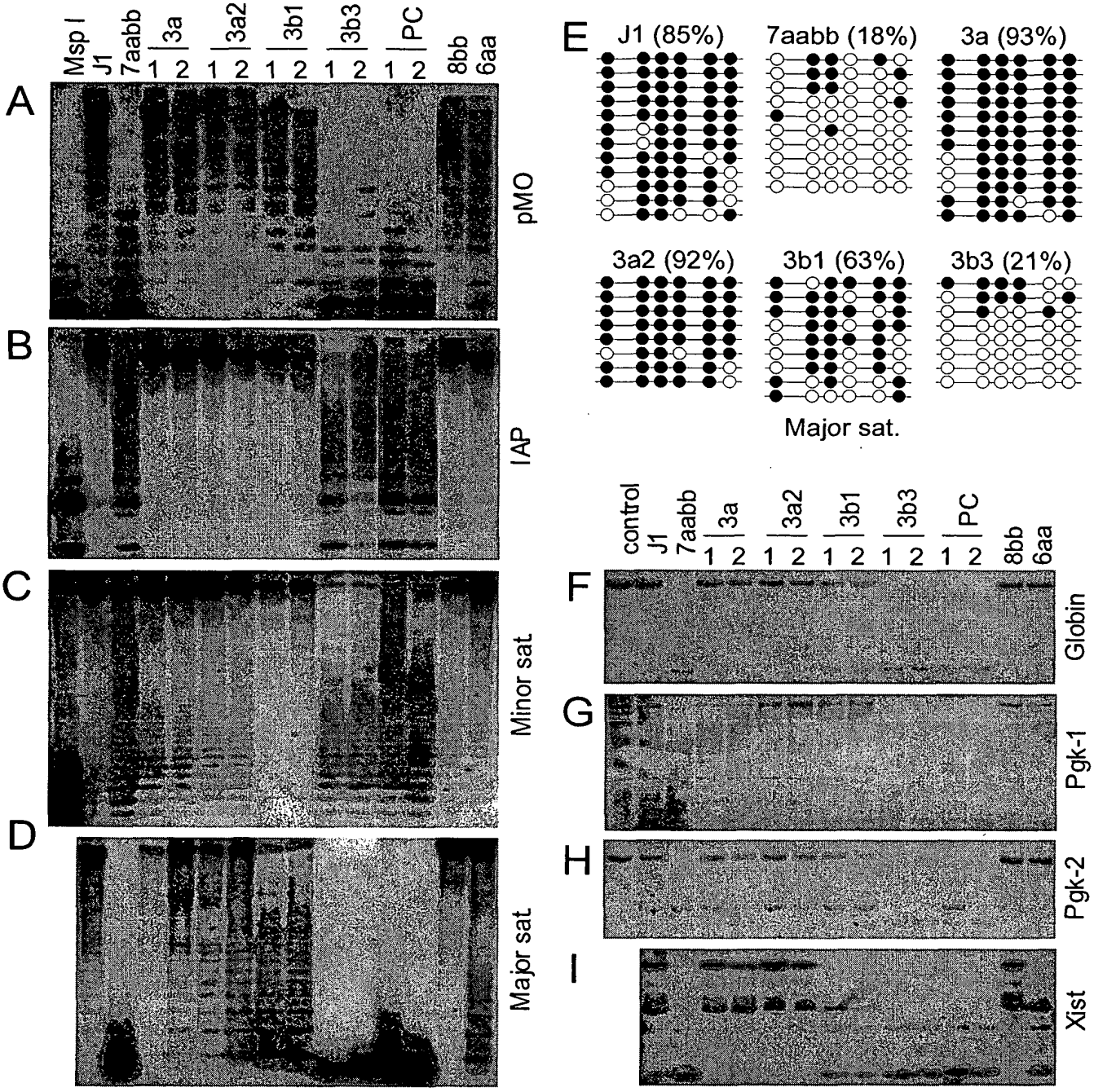

Msp I
J1
7aabb
3a
3a2
3b1
3b3
PC
8bb
6aa
1 2
A
pMO
B
IAP
C
Minor sat
D
Major sat
E
J1 (85%)
7aabb (18%)
3a (93%)
3a2 (92%)
3b1 (63%)
3b3 (21%)
Major sat.
control
F
Globin
G
Pgk-1
H
Pgk-2
I
Xist

FIG. 21

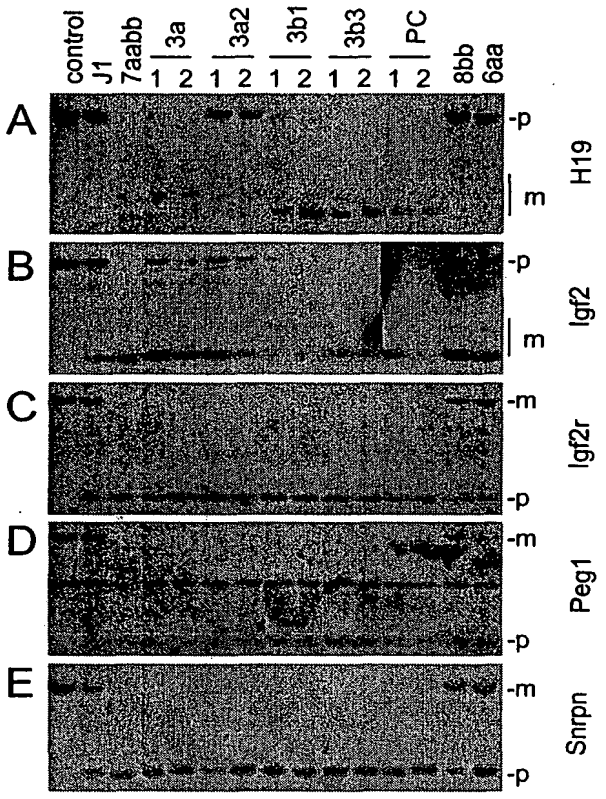

control
J1
7aabb
3a
3a2
3b1
3b3
PC
8bb
6aa
1 2 1 2 1 2 1 2 1 2
A
B
C
D
E
-p
-m
H19
Igf2
Igf2r
Peg1
Snrpn

FIG. 22

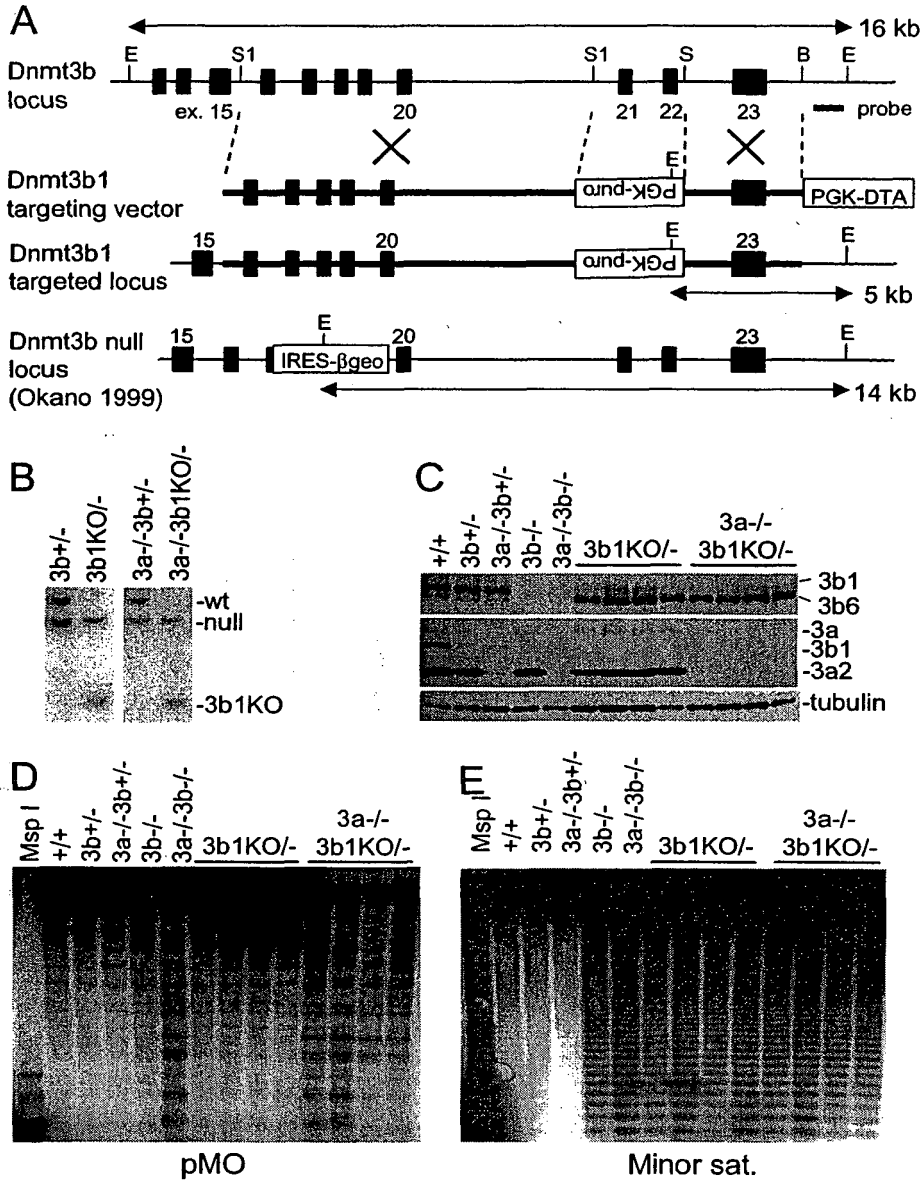

A
16 kb
Dnmt3b locus
E
S1
S1
S
B
E
ex. 15
20
21
22
23
probe
Dnmt3b1 targeting vector
PGK-puro
PGK-DTA
Dnmt3b1 targeted locus
15
20
PGK-puro
23
5 kb
Dnmt3b null locus (Okano 1999)
IRES-βgeo
14 kb
B
3b+/-
3b1KO/-
3a-/-3b+/-
3a-/-3b1KO/-
-wt
-null
-3b1KO
C
+/+
3b+/-
3a-/-3b+/-
3b-/-
3a-/-3b-/-
3b1KO/-
3a-/- 3b1KO/-
3b1
3b6
-3a
-3b1
-3a2
-tubulin
D
Msp I
+/+
3b+/-
3a-/-3b+/-
3b-/-
3a-/-3b-/-
3b1KO/-
3a-/- 3b1KO/-
pMO
E
Msp I
+/+
3b+/-
3a-/-3b+/-
3b-/-
3a-/-3b-/-
3b1KO/-
3a-/- 3b1KO/-
Minor sat.

FIG. 23

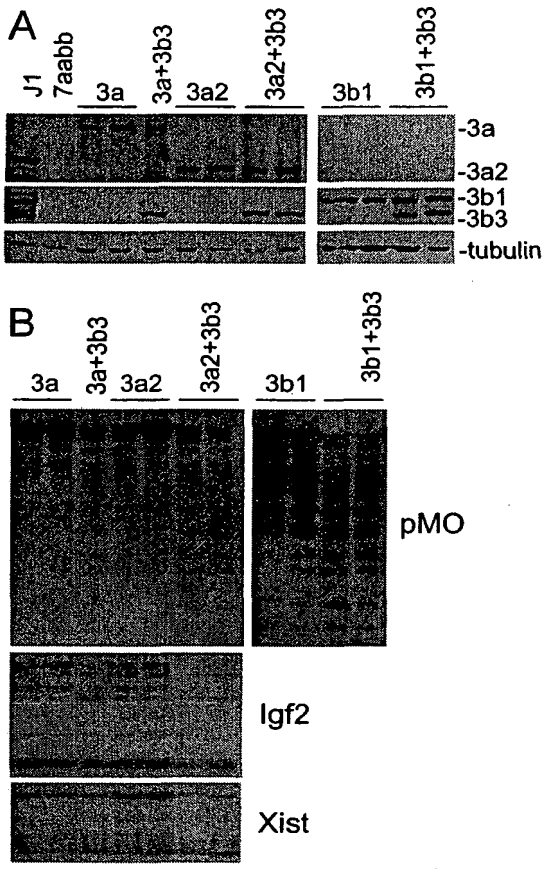

A
J1
7aabb
3a
3a+3b3
3a2
3a2+3b3
3b1
3b1+3b3
-3a
-3a2
-3b1
-3b3
-tubulin
B
3a
3a+3b3
3a2
3a2+3b3
3b1
3b1+3b3
pMO
Igf2
Xist

FIG. 24

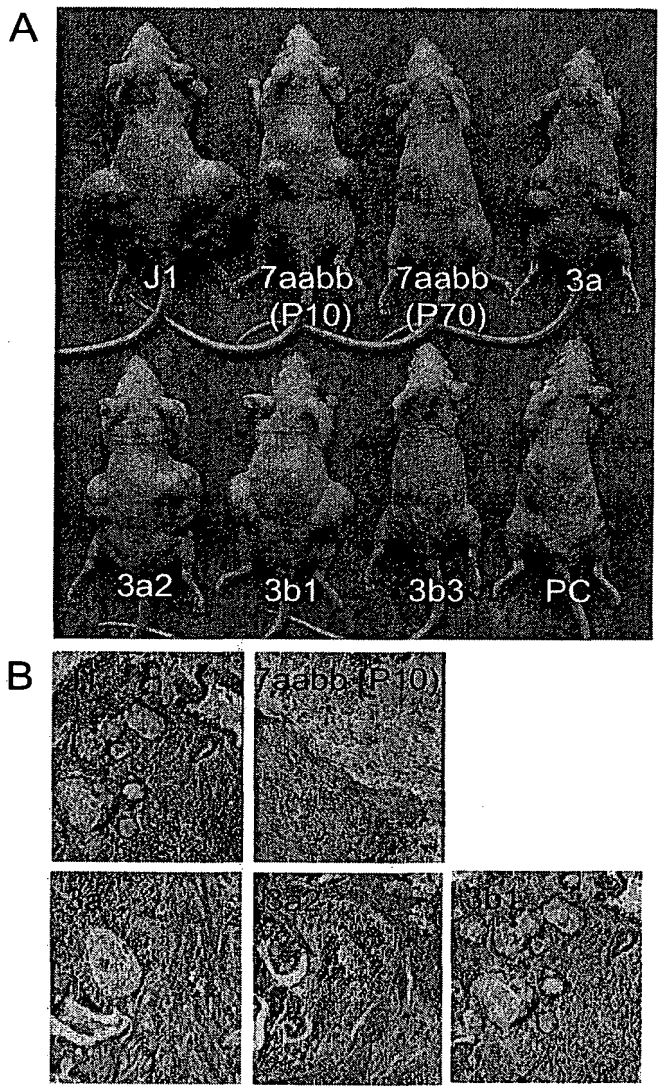


FIG. 25

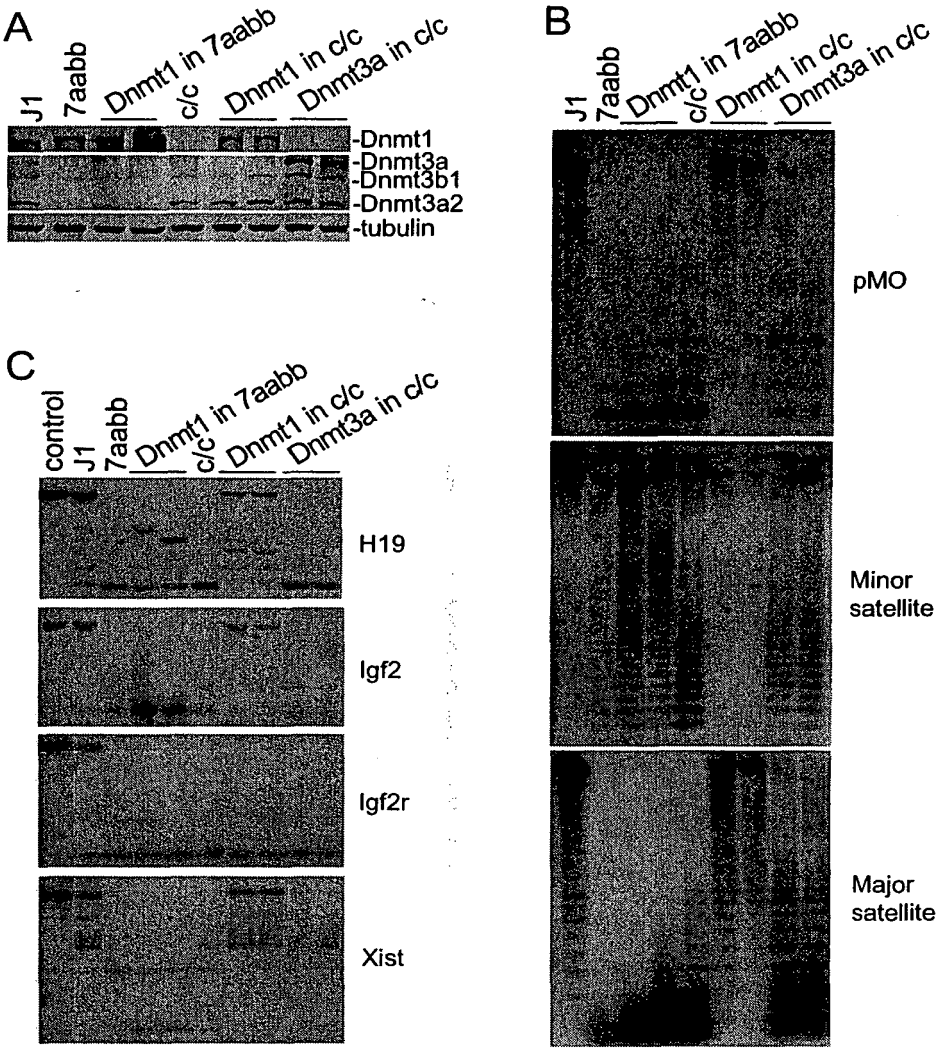

A
J1
7aabb
Dnmt1 in 7aabb c/c
Dnmt1 in c/c
Dnmt3a in c/c
-Dnmt1
-Dnmt3a
-Dnmt3b1
-Dnmt3a2
-tubulin
B
J1
7aabb
Dnmt1 in 7aabb c/c
Dnmt1 in c/c
Dnmt3a in c/c
pMO
Minor satellite
Major satellite
C
control
J1
7aabb
Dnmt1 in 7aabb c/c
Dnmt1 in c/c
Dnmt3a in c/c
H19
Igf2
Igf2r
Xist

FIG. 26

Mouse Dnmt3a2 Promoter Sequence

1 GGAGCCAGGCACCTGGGGTGTTACCTCAGTGCCTTTAGGATATTGGTTTTCCTAGCTCT
60 AGAGGGCTGATGTCATCACCCCTATTTTGCAGATGAGAAAACAGACATCTTGGGGTTAA
119 GTGGTCTGTGTCAAGGTCACCGCAATGGGATCAGGTCTTCCCCAAGCGTTCCAGCCAGA
178 TAGCGGCGGCTCCCTGCTGGGGCATTCTCCTTCAGTTCTTTGTTCTAATTCATCTTGCA
237 AACTTAATCCTGGCTAATCTTTGTAAAATACTCATTCACCTTGTTTTTCCAGAACATCT
296 GCCATGTTACAGAATATCTCCATTCAGTGCTTGACCCCAGTCCCACTACTCAGCCATTT
355 AGCTTTAGTCAAAATTGAGAGGGTGGGTGGAAGAGTTCTTTCTTCCTTCTACCTGCTTG
414 CCACCTCCAAATCGTGGTTATCTTCTGATCTCTACTGTCCTATCTCTCACCCACACCCT
473 TCATTTGATGCAGCCTTCTGCTATCTGCTTGGTGGTTTGGGTAGTTATCCACACAGGAG
532 TTTGCTTTTCAGTGATTCCCCCTTCCCCCACCCCATCTCCCCAAGTCTAGTGGAATCTA
591 TCAACTTCCTGAGAGCAGGACCAAGTGTCCATTTCTGTATCCGATGATGCTCCAGTCCT
650 CTAATGGGGGGGGGGGCGGGGCGCCAGGAGTGGCGTGTGTGCTTCTTCAAACCCAACTT
709 TAGTCCTCTACTGTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
768 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
827 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
886 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
945 NNNNNNNNNNNNNNNNNNNNNNNNNNNNACACACCTTGGGGTACTATGTCTTTGCTCAG
1004 GAATGGCATGAAATGGCCTACACTTTACCTGGTGGTTCTAGGAGAGAGACACTAGCACG
1063 TGCGTGGGAGTGTGTCTATTACTATTACATAATTGCTGAGACAGGGTTTCGTGATGTTC
1122 AGGCTGGCCTTGAACTTGTGTTAGTCAAGAATGATCTTAAATTTCTGATCTCTGGTTTC
1181 CCAAGTTCTAGGATTACAGGTGTACTTCACCACCAAAAGTTTGAACAGCTGCAGATGCC
1240 TTGGCATTGCTCTTAACGAACAGAAAATGAAACAAGCAAGCAAGACCCATTGTGACCCG
1299 GGGGACTCGGGGACTGGACGGGGAAGTTTTCAAAGTCTACTTGTGAACCACGCTTTTTA
1358 AAGCACCCCCTCCATTCACCTGTAGCGTGGCGGTGAAGTTATTGTCCTGGGGCGCCCTC
1417 AACCTGCGTGGGACACCTCCTATCCACTCACATCTGTCTTCTGACTTTGCCTAAACTAC
1476 GTTTCCGTAAACTCCGAGCCTCATCTCTAATCTGTAAACTTGCTAGCGCGCTCTCGCAC
1535 GCGCTCTTTTTTTTTTTTTTTTCCCGGAAACTCACTTTCTACAACTTTCTCCCCGGAC
1594 TCTCAGGCTGTCTGAAGCCAGCGCTCCTGTCCCACCACCGCTGCTCTGGGTGCCCCGCG
1653 GCCCGCACGCACCCTGCCTCCCTCAAGGTCCCCAACTTCCCTATGTACCCCCCCATCCC
1712 CAGAGTTGGGGGAAGGGAGCAGAGCGGGCTGTCCCATAAACCTGGCTGGAGGGGCGGGG
1771 CCCTGGGAACGGACTGGCCAGCCTCTCCCCCAGGCCCCCCGCGCCCCTCGGGCCCGGGT
1830 GAGGGGCTGGCCCAGCGCCAGCGTAGGAGGCCGGCCCCCTCCCCCCGGCCCGCGCTTAG
1889 CCAACCAGAAACTCCAGTGGGGCCCACGTGACCTGGAGTTCTAGACAAAGAAAATGTTC
1948 CCTCCCTCCCCCCCCGGCGCCCCCTCCCCCTCCCTCTGGCCCCCT***CCGCCCCCAACCCCA***
2007 ***ACGCCCCCTGCCCCTCCCCCCAGACGGGCAGCTATTTACAGAGCTTCGGGCCGGGGCTC***
2066 ***ACACCTGAGCTGTACTGCAGAGGGGCTGCACCTGGCCTTATGG***

FIG. 27

Human DNMT3A2 promoter sequence

1 GGAGCCAGGCACCTAGAGAATTGTCTCATTGTCATTAGGAGATGGTGGCGTTCCATG
60 GCCAAAGAGGGCTGATGTCATCACTCGTTTTGCAGATGAGACAACAGATTTCTTGGG
119 GGTTAAGTGACTTGTTTAAGGTCATGGTGGTGGAAACAGAACTGAAGTCCAGATCTT
178 TTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTTGCCCAGGCTGGAGTGCAGTG
237 GCATGATCTCGGCTCACTGCAACATCCGCCTCCTAAGTTGAAGCGATTCTCTTGCCT
296 CAGCCTCCCAAGTAGCTGGGATTACTGGCGCACGCCACCACGCCTGGCTAATTTTTG
355 TATTTTTAGTAGAGACAAGGTTTCACCATGTTAGTCAGGCCGGTCTCAAACTCCTGA
414 CCTCATGATCCGCCTGCCTCAGCCTTCCAAAGTGCTGGGATTATAGGCGTGAGCCAC
473 CGCGCTCGGCCAAGTCCAGATCTTCTAACAAGTGCCGCTGCCCAAATAGCCCTCTGC
532 TGTGGGGTGCATTTTCCTCCATTTCCTCAGTTCTTCCTTCTAATTCATCTTGCCAAC
591 GGCAACTAGGCTGATTTTTCCAAAATACTCATTCATCTTGTCAGAAAACCTGCGGTT
650 ATTCTTCCCTGCTACAGAATATACCCAAGGACGCACCTGAAGGCTTGCCATTACCTT
709 GCCCTGTCGTGTACTGGGAGGGTGGAGGTGGGCGAGGGTCTCCTCCCTCCCCAGCCC
768 GGCAGCTCTTGCTCATCCTACCCATCTCACCTCATTCCAAGTCCGATCCAGCCTCCA
827 GGCCCAGTCGGCTCACCTGGAACTGACCTCTGACCTCTTTTGTCATCCATGCCGCCC
886 ATTTTTTTCTACTTGGTATTTGTGGCATAGTTACCTTTACATATGTTTGTTTTACAG
945 TGATCCTTTCATATTTCTCCAAGTCTAGTGGAATCTTCAACCCCTCGAGGGCAGAGC
1004 CAACAGGGTCTATTTCTTTATCTGATCCTACAGCCAACGTAATGGAGGGCTGTGGGT
1063 GGGGACTGCGTCTGCCTTGGGGGTAGGTGCCTTTGTTCAGGAGGAGGAAGCTTGAAA
1122 TGGCGGAGGCTGCACCTGGAGGCCGCACCTGGAGGCCCCAGGAGAGGAGTCAGGTCT
1181 TCTCGATCTGCAGATGTTTGAGCCTGGGAATGAAGGAATTGCTGAACTTTCTGAAGG
1240 AGCGCCCTCGCCGCGACCAACCTTGCAAACAGGAAAATGAGAAATCCAGGGAAGGCC
1299 CAGAGTGACGCAGGGGCCCTGGGACTCGAAGCCTGACCTCCTCACGCCGCGCTTTTT
1358 GAGGCCCCCCCGCTTCTCTATTCACCTGTAGTGTGGAGGCGGGAGACCCCCCAAACA
1417 ATCCCCGATCTGGAGCGCTCCCAATGCCTGCGCGCGCCTGCTGTCACTCTCCGTCTG
1476 TGTGCTGAGTTTTCCTACAGCTTCCTGGGCCTCCTATCTGTAAGCTTTTTCTTTTTT
1535 TTTTTTTGGTTGTGCTTCAGAGAAACTCACTTTTCACAACTTTCTCCCGGCTCTCCC
1594 AGGCCGTCCGAAAGCTCCGGCTTGCTTTCGCCCGGACCCCCGGCTCCCTCCGGGCAG
1653 GCGGCTCGGGAGCAGCCCCTTCCCTCCCCTCCCGGCCCCCCGGCCCCGCGCTAATCT
1712 CTTCCAGAGCTGGGGGAGGGGCCAGGCGGTCTTCCCGAAGGCGGGGCGCTCCCTGCA
1771 GCCCGGCCTGGGCGGGCCCTGGGAACGGGCGGGGAACGGCCTCGCCCCCCGGCCCCG
1830 CGCCCCTCGGACCGGAGAAGAGGGGCTGGCCCAGCGCCAGCGTCGGAGCGCCGGCCC
1889 CCTCCCCGGGCCGCTCGCAGCCAACCAGGCCCTCCAGCGGGGCCCACGTGACCTGGA
1948 GTCCTAGACAAAGAAAATGTTCCCTCCCTCCCCCCCGCCGCCCCCCTCCCCTCCCAG
2007 TGGCCCCCT***CCGCCCCCAGCCCCATCGCCCCCTTCCCCTCCCCCAAGACGGGCAGCT***
2066 ***ACTTCCAGAGCTTCAGGGCCGCGGCTCACACCTGAGCGCGACTGCAGAGGGGCTGCA***
2125 ***CCTGGCCTTATGG***

FIG. 28

Mouse and human Dnmt3a2 promoter alignment

Top Sequence = mouse Dnmt3a2 promoter, 1858 bp (gap not counted)
Bottom Sequence = human DNMT3A2 promoter, 2065 bp

1-104 (1-105) 77% ==
1289-1338 (1475-1530) 82% ==
1518-1858 (1724-2065) 87%

```
      0     .    :    .    :    .    :    .    :    .    :
      1 GGAGCCAGGCACCTGGGGTGTTACCTCAGTGCCTTTAGGATAT  TGGTT
        |||||||||||||| | |  ||  |||| || | |||||| ||--|||
      1 GGAGCCAGGCACCTAGAGAATTGTCTCATTGTCATTAGGAGATGGTGGCG

     50     .    :    .    :    .    :    .    :    .    :
     49 TTCC TAGCTCTAGAGGGCTGATGTCATCACCCCTATTTTGCAGATGAGA
        ||||-| ||   |||||||||||||||||| | |--|||||||||||||
     51 TTCCATGGCCAAAGAGGGCTGATGTCATCACTCGT  TTTGCAGATGAGA

    100     .
     98 AAACAGA
         ||||||
     99 CAACAGA

      0     .    :    .    :    .    :    .    :    .    :
   1289 TCTTTTTTTTTTTTT  TT T  TCCCG GAAACTCAC TTTCTACAACT
        |||||||||||||||--||-|--| | |-|||||||||-||||-||||||
   1475 TCTTTTTTTTTTTTTGGTTGTGCTTCAGAGAAACTCACTTTTC ACAACT

     50     .
   1332 TTCTCCC
        |||||||
   1524 TTCTCCC

      0     .    :    .    :    .    :    .    :    .    :
   1518 GGGCCCTGGGAACGGAC TGG CCAGCCTCTCCCCCAGGCCCCCCGCGCC
        |||||||||||||||| |- ||- | ||||| ||||| ||||||--|||||
   1724 GGGCCCTGGGAACGGGCGGGGAACGGCCTCGCCCCCCGGCCCC  GCGCC

     50     .    :    .    :    .    :    .    :    .    :
   1566 CCTCGGGCCCG GGTGAGGGGCTGGCCCAGCGCCAGCGTAGGAG GCCGG
        |||||| || |-|  ||||||||||||||||||||||||| ||||-|||||
   1772 CCTCGGACCGGAGAAGAGGGGCTGGCCCAGCGCCAGCGTCGGAGCGCCGG

    100     .    :    .    :    .    :    .    :    .    :
   1614 CCCCCTCCCCCCGGCC CGCGCTTAGCCAACCAGAAACTCCAGTGGGGCC
        |||||||||| -||||-| |||--|||||||||||    |||||| ||||||
   1822 CCCCCTCCCCG GGCCGCTCGC AGCCAACCAGGCCCTCCAGCGGGGCC

    150     .    :    .    :    .    :    .    :    .    :
   1663 CACGTGACCTGGAGTTCTAGACAAAGAAAATGTTCCCTCCCTCCCCCCCG
        |||||||||||||| |||||||||||||||||||||||||||||||||||
   1869 CACGTGACCTGGAGTCCTAGACAAAGAAAATGTTCCCTCCCTCCCCCCCG

    200     .    :    .    :    .    :    .    :    .    :
   1713 GCGCCCCC TCCCCCTCCCTCTGGCCCCCTCCGCCCCCAACCCCAACGCC
         |||||||-|||||-||||  |||||||||||||||||| ||||| ||||
   1919 CCGCCCCCCCTCCCC TCCCAGTGGCCCCCTCCGCCCCCAGCCCCATCGCC

    250     .    :    .    :    .    :    .    :    .    :
   1762 CCCTGCCCCTCCCCCCAGACGGGCAGCTATTTACAGAGCTTC GGGCCGG
        |||| |||||||||| |||||||||||| || |||||||||-||||||
   1968 CCCTTCCCCTCCCCCAAGACGGGCAGCTACTTCCAGAGCTTCAGGGCCGC

    300     .    :    .    :    .    :    .    :    .
   1811 GGCTCACACCTGAGC TGTACTGCAGAGGGGCTGCACCTGGCCTTATGG
        ||||||||||||||- |-|||||||||||||||||||||||||||||||
   2018 GGCTCACACCTGAGCGCG ACTGCAGAGGGGCTGCACCTGGCCTTATGG
```

FIG. 29